The Dissertation Committee for Juan Diego Palacio Mejía Certifies that this is the approved version of the following Dissertation:

Ecology and Evolution of Cold Tolerance in *Panicum hallii*

Committee:

Thomas E. Juenger, Supervisor

Shalene Jha

David B. Lowry

Stan Roux

Randy Linder

Ecology and Evolution of Cold Tolerance in *Panicum hallii*

by

Juan Diego Palacio Mejía

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

December 2018

Dedication

To my wife, Amalia.

Acknowledgements

This dissertation is a result of an interaction of a collective of people. First, I want to express my gratitude and foremost to my advisor Dr. Thomas E. Juenger for giving me the opportunity to be part of his laboratory, for his advice, guidance, and support. I want also to acknowledge to my committee members: Dr. Shalene Jha, Dr. Stan Roux, and Dr. Randy Linder for giving advice since early stages of the research. Specially thanks to Dr. David B. Lowry, for the collections, crosses and initial steps building the natural history knowledge of *Panicum hallii*.

I want to thank the following collaborator for providing seeds and logistic help during the collections: L. González and John Reilly from the USDA Kika de la Garza Plant Materials Center. Dr. Taylor Quedensley for the collections in Texas. The collections in States of Durango and Coahuila, México, were possible thanks to the help of Dr. Socorro González, Director of the Herbarium at the Instituto Politécnico Nacional, Durango, México and the enthusiastic members of the Herbarium: Fernando Collin, Lizeth Ruacho, Jorge Tena, and Jorge Noriega. In central México, I was supported by the help of Dr. Sergio Segura and Felix Rojas from Universidad Autónoma Chapingo, Morelia, México and the Dr. Leonel González.

The field and greenhouse experiments were possible thanks the great job done by Jason Bonnette and his team: Jacob Hailing, Queen Hiers, Laxman Adhikari, Lindsay Mayer, Jill Breeden, Mathew Donahue, and Dallas Miller. Also, additional help was supported by Oscar Vargas, Ivan Vargas, and Nathan LeClear.

My fellow graduate students at Juenger lab played a key role in this work. Samsasd Razzaque for his friendship, enthusiastic help during wet lab work and data analysis. Taslima Haque for her patience, data analysis skills and time. This work would not be

possible without her commitment to help. Dr. Liz Milano and Albina Khasanova for being great office mattes.

For gathering data and assistance during the experiments I am deeply gratefully with the following undergrad students: Elena Pinaroc, Heather Yang, Maya Rao, Tina Schroyer, Beth Alger, Robert Cochran, Madison Halloway, Joshua McCauley, Roberto González, Briony Parker, Claire Mayorga, Jaclyn Heiser, Dan Carver, Shayan Bhathena, Madeline Schill, Sydney Stole, Austin Decker, Jack Donaho, Sam Parry and Jack Dwenger. Special thanks to María Villalpando who helped in different projects for around three years.

My bioinformatic skills have been one of my weakness, fortunately, my friend, Dr. Edgardo Ortiz was there to help me with his amazing ability to deal with massive molecular data. Thank you, my friend.

In the wet lab procedures, I want to say thanks to Tierney Logan and Bethany Watson for their help. Special gratitude to Galina Aglyamova who helped me with the RNA sequence library preparations.

I was very lucky to work in a laboratory with a lot of post doctorate researchers which mentored me in different aspects. First, I want to express my gratitude to Dr. Eli Meyer for sharing with me his gene expression skills. Dr. Paul Grabowski for his help with the population genetics analysis and writing process. Dr. Xiayou Weng and Dr. Eugene Shakirov, for their help during data analysis and molecular lab techniques. Dr. David Des Marais, Dr. Kyle Hernandez, Dr. Scott Schwartz and Dr. John Lovell for their help with data analysis. Special gratitude with Dr. Brandon Campitelli for his help during the experiment design and writing process. Finally, I want to say thanks to Drs. Alice McQueen, Weile Chen, Grace John, and Joseph Edwards for their suggestion in the manuscript.

I want also to thank the Genomics and Bioinformatics Service at Texas A&M University (College Station, TX), specially to R. Metz for its collaboration generating the ddRAD data. Department of Energy Joint Genome Institute, Walnut Creek, CA, for the resequence data. HudsonAlpha Institute for Biotechnology, specially Jane Grimwood for sequenced the RNA libraries. Texas Advanced Computing Center (TACC) to provide computational resources.

I want to thank to a large support network: Dr. Antonio Castilla, Will Shim, Robby Deans, Dr. Jesse Weber, Dr. Mary Lou Price, Bill Car, Andy Fang, Décio Correa, Deise Gonçalves.

I want to recognize the productive discussions with Dr. Norma Fowler and Dr. Lisa Boucher, they improve the discussion of the results.

I would like to thank my family, my mother Marta and my sister Lina for their constant support.

Special thanks to Tamra Rogers, the Graduate Coordinator for all the help from the beginning of my application to the graduate school until the graduations. She always had a solution for each problem.

This work was founded by the National Science Foundation, Plan Genome Research Program Award (IOS-0922457), Department of Energy (128996), Linda Escobar Award, and Texas EcoLab.

Lastly but not less, I want to say thanks from the deepest of my heart, to my wife Dr. Amalia Díaz, for her constant emotional and academic support.

Abstract

Ecology and Evolution of Cold Tolerance in *Panicum hallii*

Juan Diego Palacio Mejía, Ph.D.

The University of Texas at Austin, 2018

Supervisor: Thomas E. Juenger

Given that plants are sessile organism, they are continuously challenged by both biotic and abiotic stresses in natural and agricultural conditions which they are forced to acclimate or adapt. One of the major abiotic stressors affecting plant growth and development, and thereby limiting distribution of plants and crop production, is low temperature. It is estimated that crop yield is affected by cold stress in 57% of the global land area. To study the ecology and evolution of cold tolerance in plants, Hall's panicgrass (*Panicum hallii*), a perennial C_4 grass native to North America, was used to investigate cold tolerance due to its adaptation to diverse environments across its natural distribution. Three approaches were followed in this study: 1. The genetic diversity and population structure as well the response to cold stress were evaluated in samples collected across the natural distributions of the species. 2. Then, genomic regions associated to response to cold temperatures were found by QTL analysis of a RIL population. 3. Finally, a global transcriptome analysis was conducted in a set of selected phenotypes to identify candidate genes involved in cold responses. For the population genetic approach, we evaluated the genomic and morphological diversity as well as the genetic structure of the C_4 grass *Panicum hallii* using ddRAD molecular approaches. We found strong genetic and

morphological divergence between varieties and strong genetic structure between the seven populations that was strongly correlated with geological and ecological conditions. Following the genetic diversity analyses, we utilized QTL and global gene expression approaches to determine the early chilling stress transcriptome response of *P. hallii*. This analysis revealed 16 genes occurring within a major QTL interval and exhibiting expression responses. Six genes were previously associated with responses to cold stress in studies with other plant species, and five genes were identified with unknown functions. Accordingly, the results here not only aid in the discovery of the genetic mechanism that underline local adaptation but also provide a foundation to improve switchgrass yield under cold conditions.

Table of Contents

List of Tables

List of Figures

Chapter 1: Geographic patterns of genomic diversity and structure in the C_4 grass *Panicum hallii* across its natural distribution

ABSTRACT

Geographic patterns of within-species genomic diversity are shaped by evolutionary processes, life history, and historical and contemporary ecological factors. New genomic approaches can be used to understand the phylogeographic processes that have played an important role in the current distribution of genetic diversity of species. In this study, we evaluated the genomic and morphological diversity as well as the genetic structure of the C_4 grass *Panicum hallii* across its complex natural distribution in North America. We sampled extensively across the natural range of *P. hallii* in Mexico and the United States to generate ddRAD sequence data for 423 individuals from 118 localities. We used these individuals to study the divergence between the two varieties of *P. hallii*, *P. hallii* var. *filipes* and *P. hallii* var. *hallii*, and the genetic diversity and structure between populations. We also evaluated the possibility of admixture in the sympatric zone shared by both varieties. There was strong genetic and morphological divergence between the varieties with varieties, with var. *filipes* more recently derived from a var. *hallii* ancestor. We detected strong genetic structure between the seven populations that was strongly associated with geological and ecological conditions. We found evidence of admixture between the two varieties in the sympatric zone. The results of this study suggest that the current genomic and morphological diversity in *P. hallii* might be explained by diverse ecological conditions, geological history, and admixture events between the varieties.

INTRODUCTION

Geographic patterns of within-species genomic diversity are shaped by evolutionary processes, life history of species, and geological and ecological factors. Evolutionary processes generate new diversity through random mutations which are under the influence of

microevolutionary forces such as genetic drift, migration and selection (Vellend & Geber, 2005). A plant species' life history, including breeding system and dispersal mechanisms, has an influence on the genetic diversity and distribution of species (Hamrick & Godt, 1996; Ellegren & Galtier, 2016). For instance, in the grass *Brachypodium distachyon*, population structure is correlated with flowering time differences across a wide geographical range in Europe and Asia (Tyler *et al*., 2016). Current plant species distributions are the result of geological events that have had influence on soil patterns and changes in climate. For example, the soil composition of the Atlantic coastal plains in North America has been influenced by sediment deposits from the Late Cretaceous to recent deposits from the Pleistocene (Noss *et al*., 2015). Also, climatic fluctuations during the Quaternary have caused the contraction and expansion of species distribution, with strong genetic consequences (Hewitt, 2000). For example, shifts in a species' distribution result in the loss of alleles because of bottlenecks during expansion, and secondary contacts of genetic lineages differentiated in allopatry during the contractions (Hewitt, 2001, 2004). Finally, the current ecological conditions covering a species' distribution can drive genetic structure by local adaptation to these environments. Linking these processes across spatial and temporal scales can help identify the drivers of the current genetic structure of plant species populations.

Here we explored the genetic effect of different environmental conditions at the population level in Hall's panicgrass, *Panicum hallii* Vasey. This species is a perennial C_4 grass native to North America with a distribution that spans from the southeastern part of Mexico through the South-Central and Southwestern regions of the United States (Figure 1.1A). In its natural range, *P. hallii* is found across several environmental gradients such as mesic to xeric (east to west), warm to cold hardiness zones (south to north), and altitudinal, from sea level along the coastal shore of the Gulf of Mexico, Texas (Galveston, Port Aransas, and Corpus Christi) to over 2,200 meters above sea level in the Guadalupe Mountains, Texas. *P. hallii* is also found in a variety of soil and ecological conditions, including nine ecoregions as defined by Olson, *et all.,* (2001) (Figure 1.1B). These two

spatially and morphologically structured varieties, overlap in parts of their distributions: the widespread *P. hallii* Vasey var. *hallii* (hereafter var. *hallii*) and the more restricted *P. hallii* Vasey var. *filipes* (Scribn.) F. R. Waller (hereafter var. *filipes*). Variety *hallii* has a native distribution that extends from southern Colorado south into Mexico and from western Arizona to eastern Texas, and typically occurs on a variety of soil substrates, including sandy to shallow, dry, rocky, and calcareous soils. In contrast, variety *filipes* is typically found in areas of transition between clay soils and mesic depressions along the Western Gulf Coastal Grassland ecoregion and the Pine Oak Forest ecoregion at the Rio Grande Valley (Waller, 1976; Lowry *et al*., 2013). In these two ecoregions of south-central Texas, both varieties can be found in sympatric conditions. Although *P. hallii* is a highly inbred species (Lowry *et al*., 2012) and hybrids between both varieties have been obtained in a laboratory-controlled environment (Lowry *et al*., 2015), it is unknown whether these two varieties hybridize in natural conditions in their overlapping geographic range.

The two varieties of *P. halli* were described at the end of the 19th century as two different species, *P. hallii* and *P. filipes*. A century later, in a taxonomic treatment of the species of *Panicum* section *Diffusa*, the taxonomic status of these two species changed to varieties of *P. hallii,* based on panicle morphology (Waller, 1974) with clear morphological separation between the varieties, where var. *filipes* is generally larger than var. *hallii* with the exception of seed size (Waller, 1976). More recently, a set of 18 specific microsatellite markers was developed and validated for *P. hallii* (Lowry *et al*., 2012), revealing a lack of heterozygosity in the species, which suggests a self-fertilization pollination system (Lowry *et al*., 2013). Morphological and genetic analyses of 39 populations collected along a longitudinal transect from the deserts of Arizona and New Mexico through the savannas of central Texas showed geographic population structure for var. *hallii* (mean F_{ST}=0.6) with a split that occurs in western Texas (Lowry *et al*., 2013). Beyond these two studies, little is known about the drivers of genomic diversity of *P. hallii* across its native geographical range.

Panicum hallii is an model for C_4 perennial grasses and biofuel feedstocks because of its close relationship to switchgrass (*P. virgatum*), with an estimated divergence time from *P. virgatum* at 5.3 Ma (Zhang *et al*., 2011; Lovell *et al*., 2018). Switchgrass is an important perennial bioenergy crop (Robertson *et al*., 2017), and genomic tools have created opportunities to improve switchgrass performance rapidly. However, studies at the genomic level in switchgrass face substantial challenges due to its genome complexity (tetra- or octoploid series, large genome size, and high heterozygosity). *P. halli* has many traits that make it an attractive genomics model to better understand the genetics underlying important bioenergy-related traits in switchgrass (Lowry *et al*., 2013), in particular, its diploid genome, low heterozygosity, and selfing breeding system.

Here we used genetic, morphological, and geographic analyses of more than 400 samples to characterize the patterns and distribution of diversity in *P. hallii* across its natural distribution. This includes 351 newly collected samples, incorporating novel material from Mexico, West and Central Texas, and natural collections from four close relative species. We used both ddRAD, a reduced representative genotyping method, and seven discriminant morphological markers to address three major questions: 1) Is there divergence between the varieties of *P. hallii*? 2) How is the genomic and morphological diversity distributed within each variety? 3) Is there evidence of admixture between individuals from these two varieties when they are in sympatry? Answering these evolutionary questions will help us to understand how geological, ecological and life history traits can play a role in plant diversity distribution and allow us to use this diversity to explore important agronomical traits that can be used in *P. virgatum* breeding programs.

MATERIALS AND METHODS

Plant material collection

Plant material used in this study comes from two sources. First, we used *P. hallii* seed collections derived from previous studies (Lowry *et al*., 2012, 2013), which has a representative sampling of the species distribution in the Southwestern U.S. Second, we conducted new field

collections between 2012 and 2015 in Mexico and several Texas ecoregions (Western Gulf coastal grasslands, pine oak forest, East Central Texas forest, Texas Blackland prairies, Chihuahuan desert, and Edwards Plateau savanna), adding 351 new collections to the 298 previously collected plants. These new collections complement the previous collection effort and provide a representative sampling across the natural distribution of this species.

Newly collected seeds and whole plants were transported to The University of Texas at Austin greenhouse facilities for propagation. Herbarium vouchers were collected and deposited in the Billie L. Turner Plant Resources Center at The University of Texas at Austin, and leaf tissue was collected from these plants and stored in a -80°C freezer for DNA extraction. Seeds from the collected plants were harvested to be stored in a seed collection. Overall, we include 649 samples from these collections in this study (Table 1.1, Figure 1.1B).

Genotyping

High-quality DNA was extracted from the leaf tissue of each individual plant collected either in the field or in greenhouse, using a modified CTAB protocol (Allen *et al*., 2006). DNA concentration was quantified using a Qubit® 2.0 fluorometer and dsDNA BR Assay Kit (Life Technologies, Carlsbad, CA). The quality and purity of the genomic DNA was evaluated by running the samples on a 1% agarose gel for comparison to a low molecular weight ladder (New England BioLabs). All samples were normalized at 1 µg of DNA and stored at -20°C until used.

To obtain a genome-wide representation of genetic diversity at low cost, we used double-digestion restriction-associated DNA sequencing (ddRAD) (Peterson *et al*., 2012). Based on *in silico* digestions (Lepais & Weir, 2014; Mora-Márquez *et al*., 2017) of the *Panicum hallii* var. *filipes* reference genome v. 2.0 (DOE-JGI, https://phytozome.jgi.doe.gov), we chose the combination of NspI and MluCl restriction enzymes for the ddRAD method.

DNA samples were submitted in three sets to the Genomics and Bioinformatics Service at Texas A&M University (College Station, TX). Library preparation followed an in-house protocol.

Briefly, for each sample, around 1 µg of genomic DNA was digested with the restriction enzymes NspI and MluCl, and then adaptor ligation fragments were selected in a range of 375-650 bp using Pippin Prep technology (Sage, Beverly, MA). Finally, the library samples were sequenced using the Illumina HiSeq 2500 and 4000 (San Diego, CA), producing *c*. 2.5 billion of 2 x 100 or 2 x 125 pair-end raw reads.

Filtering sequencing data for nucleotide polymorphism calling and quality control

We obtained a total of ~2.6 billion raw reads for the 649 samples submitted, with a mean of ~4 million pair-end raw reads per sample (Mean= 3.961.472 ±3.265.748). Raw fastq read quality was evaluated with *FastQC* v. 0.11.3 (Andrews, 2018). Quality filtering was completed using BBTools v. 37.50 for trimming and contaminant removal following these six criteria: 1. Trimmed adapter sequence and discarding reads that were shorter than 35 bp. 2. Filtered reads with Phix viral genome and other common contaminants. 3. Trimmed reads that still had the cutsite "AATT" at the beginning. 4. Merged overlapping paired reads, and 5. Quality-trimmed reads.

The filtered reads were mapped to the *P. hallii* var. *hallii* v. 2.0 reference genome (DOE-JGI, https://phytozome.jgi.doe.gov), using bbmap v. 37.50 (Bushnell, 2014). We used the maximum likelihood statistical model in Stacks v. 1.47 (Catchen *et al*., 2011, 2013) to call SNPs, following the pipeline designed for ecology and population genomics (Rochette & Catchen, 2017). The *populations* program in Stacks was used to divide the samples into three biological groups: one set for all *P. hallii* samples, one set for var. *hallii*, and one set for var. *filipes*. We used a minimum minor allele frequency of 0.05 to process a nucleotide at a locus. We were concerned about poor mapping of short RAD reads, given the repetitive aspects of many plant genomes. As such, we filtered for paralog loci by removing markers with an excess of heterozygosity and excluding heterozygous loci with strong allele ratio deviations, using the *HDplot* method (McKinney *et al*., 2017). We removed markers with heterozygosity >6%, which was the maximum heterozygosity obtained using microsatellite markers in this species (Lowry *et al*., 2013) and the

read ratio deviation (1:4, D>4) to minimize the presence of loci that consist of repetitive regions that have been co-assembled (da Fonseca *et al*., 2016). We used paralog-finder v. 1.0 (Ortiz, 2018) to identify loci that likely contain co-assembled paralogs. In addition, we removed markers with more than 50% missing data and, subsequently, individuals with more than 80% missing data. Final VCF files were modified for different downstream population genomics analysis using vcfR v. 1.0.5 (Knaus & Grünwald, 2017).

To test for reference genome bias against var. *filipes* samples, we mapped reads from var. *filipes* to the var. *filipes* reference genome (*P. hallii* var. *filipes* v3.1 v. at DOE-JGI, http://phytozome.jgi.doe.gov/) and found that the number of SNPs is very similar and phylogenetic trees are concordant when mapping to either genome.

Species assignment via Sanger sequencing

As the taxonomic identification of species from *Panicum* section *Diffusa* is notoriously difficult (Waller, 1976), we sequenced two chloroplast and one nuclear region to aid in taxonomic confirmation. Samples of *Panicum* section *Diffusa* species collected from herbarium specimens at the Billie L. Turner Plant Resources Center (TEX), at The University of Texas at Austin, determined by specialists in the group were used as a control. A total of 31 samples were sequenced for the nuclear Internal Transcribed Spacer (ITS1) along with two new suitable chloroplast regions. The first marker, cp45676, spans the 3'end of the intergenic region between *ycf3* and *trnS* and includes the 5' portion of *trnS*. The second marker, cp1204, spans the 3' end of the intergenic region between *psbA* and *trnK* and includes the 5' portion of *trnK*. These markers were specially designed to differentiate species from the section *Diffusa* of the genus *Panicum*. The ingroup contained 17 control samples belonging to 11 species of the genus *Panicum*, section *Diffusa*. Also, six samples acted as known controls for both varieties of *P. hallii*. A sample of *Setaria viridis* was used as an outgroup for phylogenetic inference.

Approximately 570 bp of ITS1, 650 bp of cp1274, and 550 bp of cp45676 were amplified by polymerase chain reaction (PCR) using the same DNA samples that were sent for ddRAD. For ITS1, we used newly designed primers ddITS1F (5'- CCG TGA ACG TGT CAT CCA TG -3') and ddITS1R (5'- GGT CCG AGC ACC AAG GCG -3'). For cp1204: 1204F (5'- GGC TTG TAC TTT CGC GTC TC -3') and 1204R (5'- CGG TAC GAA CTT TTA TGC AAC G -3'). For cp45676: 45676F (5'- TAG GCA TAA TTC CCA ACC CA -3') and 45676R (5'- CGA ACC CTC GGT AAA CAA AA -3'). PCR mixtures were in 15 μL and were run on the MJR PTC-200 thermocycler. For ITS, we used the following conditions: 35 cycles of 95 °C denaturation for 60s, 56 °C annealing for 45s and 72 °C extension for 60s. For the cp1204 and cp45676, we used a two-step PCR with 20 cycles of 95 °C for 60s, 54 °C for 30s, reduced by 0.1 °C every cycle, and 72 °C for 60s, followed by 20 cycles of the same conditions but with a constant annealing temperature of 52 °C. PCR amplicons were visualized with SYBRSafe (Invitrogen, Eugene, OR) on 1% agarose gels and prepared for cycle sequencing by treatment with a 10:1 mixture of Shrimp Alkaline Phosphatase and Exonuclease I (both from USB; Cleveland, OH). Cleaned PCR products were submitted to The University of Texas at Austin ICMB Core Facility for cycle sequencing and visualization. The chromatograms were inspected by eye, manually trimmed for quality and aligned, then, the sequences were manually assembled, using Geneious v. 7.1.9 (Kearse *et al.*, 2012). A resulting alignment of 1,982 nucleotides was analyzed, using Bayesian phylogenetic inference as implemented in MrBayes v. 3.2.2. The resultant trees were compared, exported and edited, using FigTree v. 1.4.3 (http://tree.bio.ed.ac.uk/software/figtree/). In addition, because the species in the *Diffusa* section also differ in ploidy, the ploidy level was estimated in these samples using flow cytometry analysis (BD LSRFostesca SORP).

Population genomic structure and relatedness

We calculated heterozygosity for all samples, using VCFtools (--het) (Danecek et al. 2011). Based on preliminary analyses, we selected San Antonio Texas as a geographical landmark

corresponding the center of maximum diversity of the species. We calculated the distance of all samples to this landmark using the geosphere package (Hijmans 2017) in R. We used the R function cor.test (R Core Team, 2018) to test for correlation between heterozygosity and distance to San Antonio, TX for the var. *hallii* gene pools; the var. *filipes* gene pools lacked adequate sample sizes for the test.

Analyses of population relationships were assessed using two approaches

Phylogenetic estimation: Relationships between *P. hallii* varieties were assessed by performing maximum likelihood approach. A de novo genotyping analysis was performed using ipyrad v.0.5.1 (Eaton & Overcast, 2016) but due do the size of the dataset it could not be finished by this pipeline, therefore all resulting files from step 5 (calculation of loci consensus sequence within individuals) were imported into pyrad v.3.0.66 (Eaton, 2014) where steps 6 and 7 were run and the final matrices could be obtained. The analyses were done in the 'pairddrad' mode using a clustering threshold of 0.86, a minimum depth of 6 read pairs to form a locus within an individual, and retaining loci that were shared by at least four samples for the final matrix. This master matrix contained 640 samples from 6 species, but in order to run phylogenetic analyses, six submatrices were filtered from this master matrix. One containing only 396 samples from var. *hallii*, another containing only 27 samples from var. *filipes*, and a third one considering the 423 samples from both varieties. In the same way, three other matrices were filtered that additionally contained 5 samples of the species *Panicum capillare* to be used as an outgroup for each case. Also, the three .vcf matrices used for the PCA analyses were converted into the Phylip format using the script vcf2phylip.py (https://github.com/edgardomortiz/vcf2phylip) to conduct phylogenetic analysis.

Maximum likelihood phylogenetic estimation was performed on these 9 matrices using the software IQ-TREE v.1.5.5 (Nguyen *et al.*, 2015). The analyses were run using ascertainment bias correction under the model GTR+R+ASC since only phylogenetically informative SNPs were considered, a seed number of 123456 was specified to guarantee replicability, and branch support

was calculated using the SH-alrt test with 5000 permutations (option: –alrt 5000) (https://academic.oup.com/sysbio/article/59/3/307/1702850). Resulting topologies were checked for similarity and re-rooted using the *P. capillare* clade as outgroup.

PCA: The high-quality SNP dataset was used to identify *de novo* the optimal number of clusters and their relationship using Discriminant Analysis of Principal Component (DAPC) implemented in the Adegenet v. 2.1.1 package (Jombart, 2008) for the R software (R Core Team, 2018). DAPC centers on a discriminant function to maximize the differences between groups, minimizing variation within clusters (Jombart, 2008). To *de novo* identify the optimal number of clusters, we ran the k-means clustering (find.clusters) to assess groups, using the Bayesian information criterion (BIC); to estimate the number of PC's to be retained, we used the cross-validation test (xvalDapc) and selected the number of PCs with the lowest mean squared error, using 100 replicates. The SNP matrix of data was divided into three parts: the full set of 16,397 SNPs for all 423 samples of *P. hallii* species, one set of 16,595 SNPs for var. *hallii,* and one set of 13,167 for var. *filipes*.

Phenotypic divergence between *Panicum hallii* varieties

To explore the phenotypic divergence between var. *hallii* and var. *filipes*, a common garden was established on May 15 2014. 480 individuals from 76 different localities (ranging from 1 to 22 individuals per locality) were planted at Pickle Research Campus, The University of Texas at Austin, TX, US (Latitude 30.3838 N, longitude -97.9296 W), following a honeycomb layout with an interplant distance of 0.85 m. In total, seven phenotypic traits were chosen to evaluate divergence between the *P. hallii* varieties. During the months of July and August of 2014, one plant for each sampled locality was chosen to measure seed mass and panicle structure. For seed mass, 20 seeds per individual were weighed. Two representative mature panicles were collected and photographed to measure primary and secondary branch number and panicle length, using the Panicle Phenotyping tool P-trap (AL-Tam et al., 2013). On August 29th 2014, all samples were

tested for cold tolerance using the electrolyte leakage method (Campitelli *et al*., 2013; Pérez-Harguindeguy *et al*., 2013) in leaves after a -3 Cº frozen treatment. Early in Fall 2014, we harvested all plants and obtained above-ground biomass after drying samples for one week in a 65 ºC oven. Finally, for 368 days after planting (between May 15 2014 and May 18 2015), we measured leaf senescence every two weeks for each plant, using a qualitative scale from 1 to 4 (1 > 50% of leaf canopy green, 2 > 50% of leaf canopy senesced, 3 > 90% of leaf canopy senesced, and 4 complete canopy senescence). The average quality state was taken as the output for each plant across the year.

Admixture evaluation between *Panicum hallii* varieties

Genotypic and phenotypic analyses were conducted to explore the relationship between var. *filipes* and var. *hallii* in their region of sympatry. At the genotypic level, a new matrix of SNP was made including all var. *filipes* samples (27 individuals from 14 localities), samples from var. *hallii* in sympatry with var. *filipes* (24 individuals from 14 localities), and some var. *hallii* samples allopatric to var. *filipes* (26 individuals from 12 localities). The USEPOINFO model parameter included in the program STRUCTURE (Pritchard *et al*., 2000) was used to detect putative admixed individuals. In this approach, samples from known varieties are defined as "learning samples" (PopFlag=1) and STRUCTURE attempts to group the unknown samples (PopFlag=0) to assign the membership probability associated with each variety (Porras-Hurtado *et al*., 2013). For the phenotypic analysis, a principal component analysis (PCA) was conducted, using the phenotypic dataset obtained for the common gardens. Here, samples of var. *hallii* were divided into allopatric and sympatric samples.

RESULTS

Sampling

This study included a total of 649 individuals primarily representing *P. hallii*. The sampling included the entire geographic range of *P. hallii,* including collections from previously unsampled Mexican states and Texas counties (Figure 1.1 and Table 1.1). 28 collection localities were represented by a single individual, but most localities were represented by seeds or plants of two to 22 individuals (4.5 individuals in average). Species confirmation for undetermined samples were completed using nuclear and plastid markers. Bayesian phylogenetic trees made using the nuclear and plastid sequences clearly grouped with strong support our new collections against the control samples, confirming that the vast majority of our collected samples included three species belonging to the *Diffusa* section (Figure 1.2). While the presence of *P. hallii* in Louisiana was recently reported (Reid & Urbatsch, 2012) our sample from this state was found to be *P. diffusum,* not *P. hallii*. Further, one collection in central Texas was also assigned to *P. diffusum,* despite resembling *P. hallii* in the field. All samples collected in the central part of Mexico in the states of Mexico and Michoacan formed a clade with the *P. lepidulum* control sample, which had been collected in the same area. All samples from northern Mexico, from the states of Durango and Coahuila, clustered with *P. hallii*. Lastly, one sample in the eastern edge of the *P. hallii* range was *P. capillare*. In addition, to the phylogenic analysis, the ploidy level measured with flow cytometry showed the polyploid status of the undetermined samples, in contrast with the diploid condition of *P. hallii* (Table 1.5). We limited our subsequent analyses to 423 samples which had a strong phylogenetic assignment to *P. hallii*.

ddRAD Sequencing

We used the ddRAD approach to genotype the *P. hallii* samples. 99% percent of the raw reads passed the quality filtering criteria. These high quality reads were mapped against the *P. hallii* v. 2.1 reference genome assembly (DOE-JGI, http://phytozome.jgi.doe.gov/; Lovell *et al.*,

2018b). After genotyping, SNPs were filtered to cope with presumed paralogs and missing data at the markers and individual levels. Finally, after quality control steps, a matrix with 16,397 SNPs from 423 individuals of *P. hallii* belonging to 127 localities was obtained. In addition, individual matrices were made for each variety with 16,595 SNP for var. *hallii* (396 samples from 104 localities) and 13,167 for var. *filipes* (27 individuals from 14 localities). The high-confidence SNPs were fairly evenly distributed across the nine nuclear chromosomes of *P. hallii*.

Natural Distribution of *Panicum hallii*

Our sampling included regions where *P. hallii* has been described but for which contemporary samples are lacking. Several collection records show the presence of *P. hallii* samples in eastern Texas and in Louisiana (Figure 1.1A). However, after extensive field collection trips (16 localities), the only samples we collected were either *P. diffusum* or *P. capillare*. The absence of *P. hallii* from the East Central Texas Forest, Western Gulf coastal grasslands, and Pine Wood forest could indicate ecological constraints for establishment and persistence at the eastern edge of its distribution (Figure 1.1B). A similar situation was found at the southern part of the natural distribution, where extensive sampling was conducted in the central states of Mexico, including the states of Mexico and Michoacan. However, both morphology and molecular markers indicate that collections made in 11 localities are *P. lepidulum* (Figure 1.2). With these findings, the southern distribution of *P. hallii* seems to be restricted to the north of Mexico in the Meseta Central matorral ecoregion.

Divergence between *Panicum hallii* var. *hallii* and *Panicum hallii* var. *filipes*

Together, DAPC and phylogenetic analyses revealed substantial divergence between the *P. hallii* varieties as well as population structure within the varieties (Figure 1.3). Phylogenetic analyses confirmed that each variety of *P. hallii* forms a monophyletic clade (Figure 1.3) with a tree rooted by the outgroup method using *P. capillare*. These analyses revealed var. *hallii* as

ancestral to more recently derived var. *filipes*. Summary statistics show that var. *filipes* has greater diversity (H_e) than var. *hallii*, despite its narrow natural distribution (Table 1.1).

Population structure in *Panicum hallii* var. *hallii*

To determine the population structure of var. *hallii*, we investigated a set of 16,595 SNPs in 396 individuals from 104 localities. Using clustering methods, four clades were revealed by both the maximum likelihood tree and by DAPC. The four genetic clusters present a clear geographic pattern. There is a large cluster in the southwestern United States (Arizona, New Mexico, and West Texas), hereafter called the "West" cluster. In addition, there is a cluster including the Mexican samples and some Texas samples collected close to the Mexican border, identified as the "Tex-Mex" cluster. In central Texas, samples collected in the Edwards plateau and the Central Forest Grassland transition ecoregions are called the "Central Texas" cluster. Finally, a cluster is differentiated in South Texas, the area where both species coexist under sympatric conditions (the "sympatric" cluster). A diversity hotspot was found in southern Texas, where at least one representative of each cluster is within 100 km of San Antonio, TX. Diversity hotspots can indicate locations that were refugia during previous glacial maxima, and southern Texas has been proposed as a glacial refuge for other plants (Rebernig et al. 2010). To test whether patterns of diversity match expectations of a rapid expansion out of southern Texas following the last glacial maximum (expectation: populations farther from the refuge have lower diversity), we tested for correlations between heterozygosity and distance to San Antonio in the var. *hallii* clusters. The West cluster shows a significant negative correlation between heterozygosity and distance to San Antonio. The other three clusters also show a negative, albeit non-significant, correlation between heterozygosity and distance to San Antonio (Table 1.3).

In addition to the geographic structure of the genetic diversity in var. *hallii*, individuals collected in each locality usually group together as a monophyletic group in the phylogenetic tree (Figure 1.4A).

Population structure in *Panicum hallii* var. *filipes*

Despite the narrow distribution of var. *filipes* in South Texas, this variety shows strong population structure divided into three clusters (Figure 1.5): one genetic cluster which has a coastal distribution, another group with an inland distribution, and a third widespread group with individuals collected in both inland (Pine oak forest ecoregion) and coastal areas (Western Gulf coastal grasslands ecoregion) (Figure 1.5).

Phenotypic divergence between *Panicum hallii* var. *filipes* and *Panicum hallii* var. *hallii*

Principal component analysis (PCA) using seven phenotypic characters showed considerable divergence between var. *filipes* and var. *hallii* (Figure 1.6). In general, var. *filipes* had larger morphological features (excepting seeds) than var. *hallii* (Table 1.2). Also, var. *filipes* was more tolerant to cold stress (measured as a percentage of electrolyte leakage). Lastly, the lifespan (measured as senescence score across time) was longer for var. *filipes* than var. *hallii*.

The seven phenotypic characters also illustrate phenotypic divergence between genetically distinct geographic clusters (Figure 1.6). With PCA, we see clustering differences between geographic regions, especially in var. *filipes*, where the samples collected in the coastal and inland form two different clusters. Interestingly, the var. *hallii* samples belonging to the Sympatric group form a morphology-based cluster that is intermediate between the var. *filipes* and the rest of the var. *hallii* samples.

Genotypic and phenotypic analyses in the sympatric area

In Southwest Texas, where *var. filipes* and *var. hallii* coexist in sympatry, some samples appear intermediate between the two varieties based on genotype and phenotype information are clustering (Figure 1.6). Though *P. hallii* is a highly self-fertilizing species (Lowry *et al.*, 2012, 2013), the two varieties can be crossed in the greenhouse (Lowry *et al.*, 2015), and outcrossing in greenhouse conditions has been observed (X. Weng, 2018 pers. comm.). To determine if

intermediate samples resulted from gene flow between the two varieties, we generated a new matrix of 12,463 SNP, using ddRAD data from all samples collected in the sympatric region (27 var. *filipes* from 14 localities and 24 var. *hallii* from 14 localities) plus some var. *hallii* samples (26 individuals from 12 localities) collected in allopatric regions to serve as non-admixed controls.

We assigned the allopatric var. *hallii* samples to one population, HAL, and a random subset of var. *filipes* samples to another population, FIL. We then ran STRUCTURE (Pritchard *et al*., 2000), using the USEPOPINFO model parameter and the "HAL" and "FIL" population assignments. After 12 runs, six individuals labeled as var. *hallii* collected from four localities in the sympatric zone were assigned membership to both var. *filipes* and var. *hallii* groups under four different arrangements of var. *filipes* learning samples (Table 1.4), suggesting that these samples are admixed and there is at least periodic gene flow between var. *hallii* and var. *filipes* in sympatry.

Morphological analysis of the samples of var. *hallii* collected in the sympatry area showed intermediate phenotypic characteristics between samples of var. *filipes* and samples of var. *hallii* growing in allopatry. Phenotypic traits such as secondary branch number, biomass, tiller number, seed mass and senescence presented intermediate values between these two varieties. In addition, the PCA analysis of the phenotypic data showed that the sympatric localities of var. *hallii* formed a group between the samples of var. *filipes* and var. *hallii* in allopatry. This intermediate behavior is especially true for two (BEE.1A and UVL.1A) of the six samples that present admix compositions in the genotypic analysis.

DISCUSSION

In this study, we evaluated how geographical, ecological, geological, and life history patterns can shape the genetic diversity and genetic structure of the species in southwestern North America. As a model system we investigated the genomic and morphological diversity as well as the genetic structure of the C_4 grass *Panicum hallii* across its complex natural distribution using the most extensive sampling to date for this species. We generated ddRAD sequence data to study the

divergence between the two varieties of *P. hallii*, *P. hallii* var. *filipes* and *P. hallii* var. *hallii*, the genetic diversity and structure between populations, and tested for admixture in the sympatric zone shared by both varieties. The results show strong genetic and morphological divergence between the varieties and strong genetic structure between the seven populations, strongly associated with geological and ecological conditions. We found evidence of admixture between the two varieties in the sympatric zone. The results of this study suggest that the current genomic and morphological diversity in *P. hallii* might be explained by diverse ecological conditions, geological history, and admixture events between the varieties.

Genetic and morphological divergence between *Panicum hallii* varieties

We used ddRAD to genotype more than 400 *P. hallii* individuals from across the natural range of the species. We detected genomic divergence between the two varieties of *P. hallii*, var. *hallii* and var. *filipes* (Figure 1.3), as indicated by the phylogenic tree (Figure 1.3A) and the discriminant analysis (Figure 1.3B). This is consistent with the genetic architecture of differentiation detected with QTL (Lowry *et al.*, 2015) and morphology (Waller, 1976; Lowry *et al.*, 2013). The rooting of our phylogenetic tree suggests the *filipes* variety is derived from more ancestral var. *hallii*. The two varieties are typically found in contrasting habitats, with var. *hallii* growing in xeric habitats and var. *filipes* found in more mesic habitats. This habitat preference was observed in the field and confirmed with our collections. Indeed, Gould *et al.*, (2018) found that temperature and precipitation are key climatic variables correlated with population structure in *P. hallii*. However, we also detected morphological differentiation between the varieties, particularly in traits that may confer ecological adaptations. For instance, var. *hallii* has heavier seeds than var. *filipes* (Figure 1.7), which is a common pattern for plants that are typically exposed to drought conditions after germination in their natural habitat (Baker, 1972). We also detected differences in cold tolerance between the varieties (Figure 1.7), where var. *hallii* shows greater cold sensitivity than var. *filipes*. This was surprising, given that var. *hallii* inhabits colder hardiness zones and

higher altitudes than var. *filipes*. In addition, although *P. hallii* has been described as a perennial grass, life span evaluations show that var. *filipes* has longer life span than var. *hallii* under controlled common garden conditions (Figure 1.7, senescence). Short life span in var. *hallii* might be associated with other life history characters important for plant survival in dry environments, such as early flowering time, summer dormancy, and bigger seed mass (Lowry *et al.*, 2012, 2013). In combination, these traits may allow var. *hallii* to rapidly complete its life cycle during the short-wet seasons and avoid the stresses of the dry seasons and winters.

Across much of the range of *P. hallii*, there are strong geographic and ecological barriers to gene flow which promote and/or reinforce differentiation between the varieties. However, in southeast Texas, var. *filipes* and var. *hallii* grow in sympatry and are subject to similar climate conditions. Climate-related selection pressures promoting phenotypic differentiation may be relaxed, because var. *hallii* samples from this sympatric region are less morphologically differentiated from var. *filipes* than they are from var. *filipes* from allopatric regions (Figure 1.7). In this sympatric area, where geographic and ecological barriers to gene flow are weakened, we detect evidence of admixture between the varieties in at least six samples from four localities. Thus, despite an estimated divergence of ~1 M years, inter-mating still naturally occurs between the varieties. For instance, in the eastern locality close to Gonzalez, Texas, genetic analysis of nine samples showed that individuals from both varieties are present at the same location (one for var. *filipes* and six var. *hallii*). Two samples from that site appear to be admixed, suggesting natural crosses of these two varieties. The Gonzalez locality is particularly notable because the morphological and ecological characteristics are not typical for *P. hallii*. Morphologically, we were not able to definitively differentiate the two varieties in the field at this locality. Ecologically, this is the only locality in the Texas Blackland Prairie ecoregion where we collected both varieties of *P. hallii*, despite extensive collecting effort at other localities in the ecoregion. While there is detectable hybridization in the Gonzalez locality, samples of both varieties with no evidence of

admixture are also found in this locality, indicating that hybridization is not ubiquitous. Potential reproductive barriers between the two varieties have yet to be quantified, as they have been for other plant species (Lowry *et al*., 2008; Sobel & Chen, 2014). However, Lowry et al. (2015) did identify a two-locus Dobzhansky-Muller hybrid incompatibility causing sterility in a cross between *var. hallii* and *var. filipes*, which suggests that there is intrinsic postzygotic isolation between the two varieties.

Population structure within varieties

The two varieties of *P. hallii* are adapted to different environments, but we observe evidence of considerable differentiation within each variety as well. In a broad sense, var. *hallii* has a geographic break in genetic population structure in West Texas, coinciding with a division of Upper Cretaceous sediment to the West and Paleogene, Neogene, and Quaternary deposits to the East at around the 100º W meridian. This break isolates the West and TexMex groups from the other genetic clusters. Observation of this geological boundary imply that soil type may also be a driver of var. *hallii* population structure. To the east of this geological boundary is the Central Texas group, which is almost exclusively restricted to the Edwards Plateau and Central forest-grassland transition ecoregions, suggesting that there are either physical or ecological barriers constraining this group to these ecoregions. Within var. *filipes*, the Coastal and Inland genetic groups are found only in the Western Gulf Coastal Grasslands and Pine Oak Forest ecoregions, respectively, indicating that their differentiation is related to ecological differences between those regions. Interestingly, the third var. *filipes* group, which is found in both Western Gulf Coastal Grasslands and Pine Oak Forest ecoregions, has the most divergent morphological traits of any group and has a clear perennial life span, which indicates that the differentiation of this group may be related to evolution of life-history strategy. The relationship between population structure and ecoregions suggests that adaptive differences play a major role in shaping genetic diversity not only between varieties but also within varieties of *P. hallii*.

Phylogeography

There is fairly strong population structure within *P. hallii*, with two divergent varieties and several genetically distinct groups within each variety. Interestingly, Central Texas is a hotspot of *P. hallii* diversity as it contains individuals from all of the population genetic clusters we have detected. At least one locality representing each of the seven genetic groups is within 200 km of San Antonio, Texas (125.663 km^2). There are at least three non-exclusive reasons why this region is the center of genetic diversity for *P. hallii*: diverse ecological conditions, a complex geological history, and potential admixture between the varieties.

From the ecological perspective, Central Texas belongs to the recently proposed North American Coastal Plain (NACP) global biodiversity hotspot (Noss *et al*., 2015), characterized by its high endemism of vascular plants, especially in its grassy coastal biomes (MacRoberts *et al*., 2002; Noss, 2014). Second, this area contains seven ecoregions, defined as a land area containing a particular assembly of natural communities and species (Olson *et al*., 2001): Pine Oak Forest, Western Gulf Coastal Grasslands, Edwards Plateau Savanna, Chihuahua Desert, Central and Southern Mixed Grasslands, Central Forest Grassland Transition and Texas Blackland Prairies (Figure 1.2). Third, in addition to this ecoregion diversity, at the plant diversity level, a mix of floristic provinces coincide in Central and South Texas, including taxa from North American Prairies Province, Sonora Province, and Coastal Plain Floristic Province, where around 20% of the endemic plants are grasses (Sorrie & Weakley, 2001). Also, the area around San Antonio, Texas, contains five grassland community types (Fayette, Upper Coastal, Coastal, Blackland, and San Antonio Prairie) characterized by seven major soil associations, (Diamond and Smeins, 1985). Taken together, these ecological factors likely contribute to the high diversity of *P. hallii* in this narrow area in Central Texas and the Rio Grande Valley.

Geological history has also impacted the diversity in this region. It is well known that the Quaternary climatic fluctuations in the form of multiple contraction-expansion cycles of species distribution shaped the contemporary genetic composition of the biodiversity (Hewitt, 2000, 2004;

Razgour *et al*., 2013). During the glacial cycles of the Pleistocene, South Texas was a glacial refugium belt for boreal displaced vegetation from the north boreal latitudes (Holmgren *et al*., 2007) and for lowland taxa during the flooding due to sea level fluctuations (Noss *et al*., 2015). It served as a refugium for a variety of plants and animals (Elias, 1992; Swenson & Howard, 2005; Majure *et al*., 2012; Bryson *et al*., 2014; Seal *et al*., 2015), promoting species richness and genetic diversity (Swenson & Howard, 2005). A large-scale aridification of the southwestern parts of North America started during the Holocene at the end of the last glacial maximum (~21,000 ybp), allowing the range expansion of drought-adapted genotypes from the refuge areas to the new arid regions (Rebernig *et al*., 2010) following the Northwest outward-direction, post-glacial expansion route proposed by Swenson and Howard (2005) for Central Texas refugia. This possible Holocene expansion might explain the lack of diversity in the West and the TexMex group of var. *hallii* as these groups may be a product of a population bottleneck. In fact, there is a significant negative correlation between diversity (heterozygosity) and distance from southern Texas (San Antonio) for the var. *hallii* West cluster, consistent with a population bottleneck(s) during a post-glacial expansion (Table 1.3). A similar pattern has been found in blackfoot daisy (*Melampodium leucanthum*, Asteraceae)(Stuessy *et al*., 2004; Rebernig *et al*., 2010) and members of the *Humifusa* clade of *Opuntia* genus (Majure *et al*., 2012).

Finally, diversity may be high in Central Texas due to hybridization between the varieties. During the Pleistocene, there were recurrent glacial cycles and the species range likely contracted and expanded multiple times. These successive contraction-expansion events led to the formation of contact zones, hybrid zones, and phylogeographic break areas in North America (Suture zones)(Remington, 1968; Hewitt, 2001), one of them in Central Texas (Swenson & Howard, 2004, 2005), which corresponds to the same region of *P. hallii* genetic diversity in this work. Thus, the high diversity may be due to a history of secondary contact at this location between groups that diverged in allopatry during previous glacial cycles. More extensive sampling in the sympatric

zone for both varieties and the use of plastid genomic information may be helpful in addressing this question in future studies.

Panicum hallii distribution

The results from our field collection trips suggested that the natural distribution of *P. hallii* is narrower than previously considered at the southern and eastern edges of the species distribution. Previous literature and collection records indicate that the natural distribution of *P. hallii* spans from the Southwest United States (Oklahoma, Colorado, Texas, New Mexico, and Arizona) through Mexico down to the southern Mexican state of Oaxaca (Gould, 1975; Waller, 1976; Herrera & Pámanes, 2006; Estrada & Villareal, 2010; Herrera-Arrieta & Cortés-Ortiz, 2010; Sánchez-Ken, 2011; Valdés-Reyna *et al*., 2015). However, our collection efforts in the southern portion of this reported range, in the Mexican states of Mexico and Michoacán, failed to find *P. hallii*. The plants we found that most resemble *P. hallii* have morphological characteristics not found in *P. hallii* samples collected in the United States, and molecular markers indicate that these plants correspond to *P. lepidulum* (Figure 1.2). While our sampling was not exhaustive, we focused our efforts in areas where herbarium samples had been previously collected, and the absence of *P. hallii* in these Mexican states suggests that the southern boundary of the distribution of *P. hallii* is currently the Meseta Central Matorral and Chihuahua Desert ecoregions in the Northern Mexican states.

At the eastern part of the distribution, several collection records indicate that *P. hallii* grows in eastern Texas close to the Louisiana border and even in northwest Louisiana (Reid & Urbatsch, 2012). However, the samples that most resemble *P. hallii* are, in fact, *P. diffusum* and *P. capillare*. We also made at least 14 collection trips east of the Edwards Plateau ecoregion in central Texas without finding *P. hallii*, suggesting that the eastern distribution of *P. hallii* is restricted to the Texas Blackland Prairies and East Central Texas forest ecoregions (Figure 1.1).

Overall, var. *hallii* is associated primarily with six ecoregions: the Meseta Central Matorral in the South, the Chihuahuan Desert and the Arizona Mountain Forest ecoregion to the West, the Western Short Grasslands and Central Forest Grasslands Transition to the North, and finally, the Central and Southern Mixed Grasslands and Edwards Plateau Savanna to the East. The distribution of var. *filipes* shows it is restricted to the Western Gulf Coastal Grasslands and Pine Oak Forest ecoregions, where var. *hallii* was also collected but was much less abundant.

Lessons for diversity studies in morphologically and ecologically diverse systems

One challenge we encountered while trying to sample diversity from across the range of *P. hallii* was to correctly distinguish *P. hallii* from closely related species. *P. hallii* grows in a variety of habitats, and its morphology in natural and controlled settings can differ substantially. We collected 45 samples from the originally reported range of *P. hallii* that were members of other species in *Panicum* section *Diffusum*. There are many research systems that share this characteristic: ecologically and morphologically diverse taxa that are difficult to distinguish from closely related taxa in natural setting (Rebernig *et al*., 2010; Tyler *et al*., 2016). Sampling the breadth of genetic, morphological, and ecological diversity is important, but can lead to the collection of non-target taxa. While including all collected samples in sequencing-based genotyping methods (ex: ddRAD, GBS) can resolve the relationships among target and non-target taxa, that may not be the most efficient use of sequencing, computing, or analysis resources. For instance, sequencing effort of the target taxa is decreased when non-target taxa are included. In addition, in our experience, methods for identifying SNPs and calling genotypes in non-model species can be sensitive to phylogenetic diversity (Mastretta-Yanes *et al*., 2015; Shafer *et al*., 2017), and starting strictly with the target taxa reduces the filtering steps necessary to produce a useful set of SNPs for subsequent analyses.

Therefore, we recommend that, if possible, samples be checked in a controlled environment to identify samples that do not match the standard description of the taxon, though we recognize

this is not possible for many systems. Then a quick method should be used to verify that suspicious samples are within the intended taxonomic level for the study. For this study, Sanger sequencing of nuclear and chloroplast loci were fast and provided the resolution required to differentiate species. Finally, the next-generation sequencing (NGS) efforts should focus on the verified samples. While the initial steps take time and may delay NGS efforts, they are a valuable investment as they improve the sequencing results and can save substantial time during bioinformatic analysis.

In conclusion, the genetic and morphological evaluation in this study represents the most extensive sampling of the *P. hallii* varieties. This evaluation made it possible to distinguish the genetic structure at different biological levels and to infer the direction of evolution of the *hallii* and *filipes* varieties. At the species level, considerable divergence was found between both varieties using molecular and morphological data. Within varieties, it was possible to characterize population structure across the diverse geological and ecological native range of *P. hallii*. Within the sympatric zone, this set of data allowed us to identify putative admixed individuals between var. *filipes* and var. *hallii*. Taken together, these data illustrate the geographical distribution of genetic diversity in *P. hallii*. A practical application of these results could be the creation of a core collection of these samples to streamline screening for particular traits of interest. Future work could include increased sampling efforts for var. *filipes*, especially in the Rio Grande Valley and Mexico. In addition, to confirm admixture between the varieties, the plastome genome from samples collected in the sympatric area can be explored. In addition, studies of local adaptation could be conducted by crossing members of different populations (Lowry *et al*., 2015; Milano *et al*., 2016), or by reciprocal transplants across ecoregions to identify the traits that underlie ecotype formation (Hereford, 2009).

TABLES

Table 1.1. List of localities. Number of samples (N), expected heterozygosity (He).

Loc ID	Species	Region	Country	State	County	Altitude m.a.s.l.	Latitude	Longitude	N	*He*
HAR	*Panicum capillare*		USA	Texas	Harris	6	29.7197	-94.9502	5	
DSO	*Panicum diffusum*		USA	Louisiana	De Soto	51	32.3062	-93.8076	1	
ENR	*Panicum diffusum*		USA	Texas	Llano	461.3	30.5085	-98.8246	4	
ARA	*Panicum hallii var. filipes*	Coastal & inland	USA	Texas	Aransas	2	28.2237	-96.987	1	0.044
AWR	*Panicum hallii var. filipes*	Coastal	USA	Texas	Aransas	2.7	28.2851	-96.9457	2	0.076
CAM	*Panicum hallii var. filipes*	Coastal & inland	USA	Texas	Cameron	5	26.0007	-97.2691	3	0.039
DIM	*Panicum hallii var. filipes*	Inland	USA	Texas	Dimmit	162	28.5701	-99.5245	1	0.025
EPS	*Panicum hallii var. filipes*	Coastal & inland	USA	Texas	Maverick	248	28.4826	-100.21	1	0.041
FIL	*Panicum hallii var. filipes*	Coastal & inland	USA	Texas	Nueces	3	27.6496	-97.4039	7	0.059
LAR	*Panicum hallii var. filipes*	Coastal & inland	USA	Texas	Webb	160	27.4773	-99.2092	3	0.063
NDP	*Panicum hallii var. filipes*	Coastal	USA	Texas	San Patricio	1	27.9117	-97.6076	3	0.052
PAB	*Panicum hallii var. filipes*	Coastal	USA	Texas	Cameron	9	26.0199	-97.4745	1	0.054
ZAF	*Panicum hallii var. filipes*	Inland	USA	Texas	Zapata	97	26.5759	-99.1091	1	0.04
ZAP	*Panicum hallii var. filipes*	Coastal	USA	Texas	Zapata	116	27.021	-99.1251	3	0.043
GNF	*Panicum hallii var. filipes*	Coastal	USA	Texas	Dewitt	76.5	29.3119	-97.3001	2	0.038
MCM	*Panicum hallii var. filipes*	Coastal & inland	USA	Texas	McMullen	100	28.1519	-98.6385	1	
NUE	*Panicum hallii var. filipes*	Coastal	USA	Texas	Nueces	15	27.5684	-97.7956	1	

Table 1.1 (Continue)

Loc ID	Species	Region	Country	State	County	Altitude m.a.s.l.	Latitude	Longitude	N	*He*
ATC	*Panicum hallii var. filipes*	Inland	USA	Texas	Atascosa	786	28.824	-98.6932	5	0.02
ATC	*Panicum hallii var. hallii*	Sympatric	USA	Texas	Atascosa	786	28.824	-98.6932	5	0.016
RAR	*Panicum hallii var. hallii*	Sympatric	USA	Texas	Live Oak	83	28.607	-98.222	5	0.034
RAR	*Panicum hallii var. filipes*	Inland	USA	Texas	Live Oak	83	28.607	-98.222	5	0.024
ABI	*Panicum hallii var. hallii*	Central Texas	USA	Texas	Coleman	540	31.8599	-99.4422	3	0.006
ALI	*Panicum hallii var. hallii*	West	USA	Texas	Potter	961	35.5739	-101.7011	7	0.004
ARB	*Panicum hallii var. hallii*	Central Texas	USA	Oklahoma	Murray	305	34.4298	-97.1467	4	0.007
ART	*Panicum hallii var. hallii*	West	USA	New Mexico	Chaves	1441	32.8567	-104.9316	10	0.004
ATA	*Panicum hallii var. hallii*	Tex-Mex	USA	Texas	Atascosa	113	28.7897	-98.5413	1	0.015
ATO	*Panicum hallii var. hallii*	Tex-Mex	Mexico	Durango	Mezquital	1869	23.693	-104.3961	5	0.027
BAN	*Panicum hallii var. hallii*	Central Texas	USA	Texas	Bandera	679	29.8473	-99.5734	5	0.008
BBD	*Panicum hallii var. hallii*	Tex-Mex	USA	Texas	Brewster	780	29.6235	-103.1162	4	0.014
BBF	*Panicum hallii var. hallii*	Tex-Mex	USA	Texas	Brewster	1244.4	29.2736	-103.3375	3	0.011
BEE	*Panicum hallii var. hallii*	Sympatric	USA	Texas	Bee	83	28.3828	-97.7792	4	0.034
BEL	*Panicum hallii var. hallii*	Central Texas	USA	Texas	Bell	240	30.9183	-97.6581	5	0.006
BEX	*Panicum hallii var. hallii*	Central Texas	USA	Texas	Bexar	341	29.7214	-98.3542	5	0.005
BME	*Panicum hallii var. hallii*	West	USA	Oklahoma	Cimarron	1319	36.8753	-102.8874	6	0.005
BOS	*Panicum hallii var. hallii*	West	USA	Texas	Bosque	204	32.128	-97.5596	5	0.006
BSP	*Panicum hallii var. hallii*	Sympatric	USA	Texas	Brown	440	31.8616	-99.0258	1	0.015
BUR	*Panicum hallii var. hallii*	Sympatric	USA	Texas	Burnet	248	30.5468	-98.1562	5	0.013
BUS	*Panicum hallii var. hallii*	West	USA	Texas	Potter	1119	35.2649	-102.0605	8	0.004

Table 1.1 (Continue)

Loc ID	Species	Region	Country	State	County	Altitude m.a.s.l.	Latitude	Longitude	N	*He*
CBS	*Panicum hallii var. hallii*	Central Texas	USA	Texas	San Saba	346.4	31.0574	-98.4839	5	0.008
CCG	*Panicum hallii var. hallii*	West	USA	Texas	Briscoe	754	34.4468	-101.0823	6	0.006
CCL	*Panicum hallii var. hallii*	West	USA	Texas	Briscoe	742	34.4441	-101.0751	5	0.006
CCO	*Panicum hallii var. hallii*	West	USA	Texas	Briscoe	749	34.4379	-101.0561	2	0.005
CDM	*Panicum hallii var. hallii*	Tex-Mex	Mexico	Coahuila	Parras	1262	25.5083	-102.5699	5	0.016
CHA	*Panicum hallii var. hallii*	West	USA	Texas	Oldham	1069	35.6052	-102.2966	8	0.004
CKR	*Panicum hallii var. hallii*	West	USA	Arizona	Maricopa	1817	34.2342	-112.3154	10	0.004
CLA	*Panicum hallii var. hallii*	West	USA	New Mexico	Union	1463	36.6364	-103.0375	1	0.004
CLD	*Panicum hallii var. hallii*	West	USA	Texas	Armstrong	956	34.7984	-101.4363	7	0.004
CNF	*Panicum hallii var. hallii*	West	USA	Arizona	Santa Cruz	1550	31.8021	-110.9152	4	0.004
COL	*Panicum hallii var. hallii*	Sympatric	USA	Texas	Coleman	474	31.8063	-99.2348	5	0.013
COM	*Panicum hallii var. hallii*	Tex-Mex	USA	Texas	Val Verde	485	29.7018	-101.2077	2	0.013
COR	*Panicum hallii var. hallii*	Tex-Mex	USA	Arizona	Maricopa	1110	34.2862	-112.179	17	0.01
CRC	*Panicum hallii var. hallii*	West	USA	Texas	Oldham	985	35.5298	-102.2651	8	0.005
CRS	*Panicum hallii var. hallii*	West	USA	Texas	Briscoe	740.6	34.4374	-101.0568	4	0.005
DOF	*Panicum hallii var. hallii*	Tex-Mex	USA	Texas	Val Verde	417	29.8946	-100.9946	7	0.017
DRB	*Panicum hallii var. hallii*	Tex-Mex	USA	Texas	Val Verde	477	29.9412	-100.9706	5	0.015
DRS	*Panicum hallii var. hallii*	Tex-Mex	USA	Texas	Val Verde	457.4	29.9051	-101.0032	5	0.023
DUV	*Panicum hallii var. hallii*	Sympatric	USA	Texas	Duval	139	27.9141	-98.7002	3	0.012
EDW	*Panicum hallii var. hallii*	Central Texas	USA	Texas	Edwards	628	29.8263	-100.4664	5	0.006
ELG	*Panicum hallii var. hallii*	West	USA	Arizona	Santa Cruz	1531	31.6009	-110.5585	10	0.005

Table 1.1 (Continue)

Loc ID	Species	Region	Country	State	County	Altitude m.a.s.l.	Latitude	Longitude	N	*He*
ESE	*Panicum hallii var. hallii*	Central Texas	USA	Texas	Kimble	642	30.3482	-99.5894	3	0.01
GEO	*Panicum hallii var. hallii*	West	USA	New Mexico	Harding	1780	36.0429	-104.3092	1	0.006
GIL	*Panicum hallii var. hallii*	West	USA	New Mexico	Grant	2101	32.8273	-108.057	25	0.003
GMO	*Panicum hallii var. hallii*	West	USA	Texas	Culberson	2257.5	31.9018	-104.8391	6	0.006
GRA	*Panicum hallii var. hallii*	West	USA	Texas	Kerr	642	29.9149	-99.2431	1	0.016
HAL	*Panicum hallii var. hallii*	Central Texas	USA	Texas	Travis	244	30.185	-97.874	2	0.007
HAY	*Panicum hallii var. hallii*	Central Texas	USA	Texas	Hays	387	29.9653	-98.1804	6	0.004
HWY	*Panicum hallii var. hallii*	West	USA	Texas	Briscoe	931	34.4691	-101.1103	5	0.004
JHC	*Panicum hallii var. hallii*	Tex-Mex	USA	Texas	Blanco	361	30.2073	-98.307	4	0.02
JWE	*Panicum hallii var. hallii*	Sympatric	USA	Texas	Jim Wells	169	27.894	-98.0055	1	0.014
KEN	*Panicum hallii var. hallii*	Tex-Mex	USA	Texas	Kendall	478	29.8593	-98.694	5	0.009
KER	*Panicum hallii var. hallii*	Central Texas	USA	Texas	Kerr	806	30.0397	-99.3973	5	0.009
KNT	*Panicum hallii var. hallii*	West	USA	Texas	Culberson	1326	31.0499	-104.212	3	0.006
LBY	*Panicum hallii var. hallii*	West	USA	New Mexico	Guadalupe	1450	34.945	-104.6406	3	0.005
LMA	*Panicum hallii var. hallii*	Central Texas	USA	Texas	Real	678	30.0164	-99.7314	4	0.007
LME	*Panicum hallii var. hallii*	West	USA	Texas	Hutchinson	918	35.7073	-101.5465	7	0.006
LOS	*Panicum hallii var. hallii*	West	Mexico	Coahuila		1550	25.4619	-101.0239	3	0.01
MAP	*Panicum hallii var. hallii*	Tex-Mex	Mexico	Durango	Mapimi	1563	25.872	-104.1051	5	0.014
MCR	*Panicum hallii var. hallii*	West	USA	Texas	Jeff Davis	1865	30.6829	-104.1348	8	0.007
MEZ	*Panicum hallii var. hallii*	Tex-Mex	Mexico	Durango	Mezquital	1600	23.4945	-104.4324	5	0.01
MKF	*Panicum hallii var. hallii*	Central Texas	USA	Texas	Travis	174	30.1777	-97.7251	2	0.044

Table 1.1 (Continue)

Loc ID	Species	Region	Country	State	County	Altitude m.a.s.l.	Latitude	Longitude	N	*He*
NDD	*Panicum hallii var. hallii*	Tex-Mex	Mexico	Durango	Nombre de Dios	1866	23.8979	-104.335	5	0.013
NMX	*Panicum hallii var. hallii*	West	USA	New Mexico	Dona Ana	1656	32.4299	-106.5815	2	0.005
NOH	*Panicum hallii var. hallii*		USA	Texas	Willacy	4.57	26.4949	-97.5453	1	0.021
PAJ	*Panicum hallii var. hallii*	Sympatric	USA	Texas	Kimble	532	30.4854	-99.7376	1	0.019
PCR	*Panicum hallii var. hallii*	West	USA	Arizona	Cochise	1652	31.9802	-109.3634	1	0.004
PDB	*Panicum hallii var. hallii*	West	USA	Texas	Randall	870	34.9287	-101.6366	9	0.004
PDF	*Panicum hallii var. hallii*	West	USA	Texas	Randall	864	34.9437	-101.6613	8	0.003
PDK	*Panicum hallii var. hallii*	West	USA	Texas	Randall	905	34.9476	-101.689	1	0.004
PDL	*Panicum hallii var. hallii*	West	USA	Texas	Randall	892	34.9476	-101.689	2	0.003
PDS	*Panicum hallii var. hallii*	West	USA	Texas	Randall	860	34.9395	-101.6528	2	0.004
PDT	*Panicum hallii var. hallii*	West	USA	Texas	Randall	1040	34.9391	-101.6305	8	0.004
PFL	*Panicum hallii var. hallii*	Central Texas	USA	Texas	Travis	200	30.4506	-97.6386	2	0.006
PIN	*Panicum hallii var. hallii*	West	USA	Arizona	Gila	1642	34.3161	-111.4006	3	0.003
PIS	*Panicum hallii var. hallii*	Sympatric	USA	Texas	Cameron	5	26.0173	-97.2735	1	0.009
POI	*Panicum hallii var. hallii*	Sympatric	USA	Texas	Cameron	5	26.0643	-97.2375	2	0.012
PRT	*Panicum hallii var. hallii*	West	USA	Arizona	Cochise	1478	31.9159	-109.1547	8	0.003
REA	*Panicum hallii var. hallii*	Central Texas	USA	Texas	Real	599	29.7683	-99.8171	5	0.008
SAN	*Panicum hallii var. hallii*	Tex-Mex	USA	Texas	Terrel	981	30.1346	-102.5633	3	0.007
SAP	*Panicum hallii var. hallii*	Tex-Mex	Mexico	Durango	Lerdo	1476	25.4331	-103.7161	5	0.016
SCG	*Panicum hallii var. hallii*	Tex-Mex	USA	Texas	Val Verde	421.4	29.6942	-101.3221	5	0.014
SEM	*Panicum hallii var. hallii*	Tex-Mex	USA	Texas	Val Verde	425	29.681	-101.3099	4	0.016

Table 1.1 (Continue)

Loc ID	Species	Region	Country	State	County	Altitude m.a.s.l.	Latitude	Longitude	N	*He*
SEP	*Panicum hallii var. hallii*	Central Texas	USA	Texas	Travis	291.9	30.4013	-97.797	7	0.018
SEV	*Panicum hallii var. hallii*	West	USA	New Mexico	Socorro	1544	34.3317	-106.9735	24	0.003
SLR	*Panicum hallii var. hallii*	Sympatric	USA	Texas	Kimble	538	30.4453	-99.7974	1	0.01
SLS	*Panicum hallii var. hallii*	Tex-Mex	USA	Texas	Burnet	391	30.9913	-98.1008	4	0.012
SMF	*Panicum hallii var. hallii*	Central Texas	USA	Texas	Lampasas	476	31.3052	-98.4387	1	0.007
SMT	*Panicum hallii var. hallii*	Tex-Mex	USA	Texas	Mills	488	31.5616	-98.5992	3	0.016
SNO	*Panicum hallii var. hallii*	West	USA	Arizona	Pima	1567	31.8005	-110.7035	8	0.004
SOM	*Panicum hallii var. hallii*	Central Texas	USA	Texas	Somervell	191	32.2756	-97.6016	5	0.007
SPE	*Panicum hallii var. hallii*	West	USA	Texas	Crockett	744	30.6664	-100.9707	23	0.005
SPJ	*Panicum hallii var. hallii*	Central Texas	USA	Texas	Kimble	602	30.47	-99.7308	5	0.006
SPR	*Panicum hallii var. hallii*	West	USA	Texas	Jeff Davis	1900	30.6723	-104.1277	1	0.006
SRG	*Panicum hallii var. hallii*	Tex-Mex	USA	Texas	Val Verde	380.5	29.6608	-101.3152	8	0.018
STW	*Panicum hallii var. hallii*	Central Texas	USA	Texas	Palo Pinto	332	32.6093	-98.4896	8	0.006
SVN	*Panicum hallii var. hallii*	West	USA	New Mexico	Guadalupe	1551	34.8311	-104.8274	9	0.004
THC	*Panicum hallii var. hallii*	Central Texas	USA	Texas	Travis	289	30.2419	-98.0091	5	0.014
TLA	*Panicum hallii var. hallii*	Tex-Mex	Mexico	Durango	Tlahualilo	1114	26.1281	-103.4992	5	0.012
TNK	*Panicum hallii var. hallii*	West	USA	New Mexico	Guadalupe	1398	34.9369	-104.694	4	0.004
TWC	*Panicum hallii var. hallii*	Central Texas	USA	Texas	Travis/Williamson	360	30.6222	-97.9719	5	0.004
UVA	*Panicum hallii var. hallii*	Sympatric	USA	Texas	Uvalde	304	29.2415	-100.0972	1	0.014
UVL	*Panicum hallii var. hallii*	Sympatric	USA	Texas	Uvalde	550	29.5983	-100.0253	2	0.037
WBW	*Panicum hallii var. hallii*	Sympatric	USA	Texas	Kimble	597	30.4329	-99.7958	3	0.045

Table 1.1 (Continue)

Loc ID	Species	Region	Country	State	County	Altitude m.a.s.l.	Latitude	Longitude	N	*He*
ARR	*Panicum hallii var. hallii*	West	USA	Arizona	Santa Cruz	1460	31.6134	-110.5063	2	
CMR	*Panicum hallii var. hallii*	Sympatric	USA	Texas	Maverick	219	28.5835	-100.1117	2	
GNH	*Panicum hallii var. hallii*	Sympatric	USA	Texas	Dewitt	76.5	29.3119	-97.3001	2	0.029
HAM	*Panicum hallii var. hallii*	Tex-Mex	USA	Texas	Stonewall	505	33.0098	-100.1813	2	
HUE	*Panicum hallii var. hallii*	Sympatric	USA	Texas	Gillespie	550	30.2255	-99.0156	1	
KIC	*Panicum hallii var. hallii*	Tex-Mex	USA	Texas	Kinney	561	29.6185	-100.443	1	
MAV	*Panicum hallii var. hallii*	Sympatric	USA	Texas	Maverick	215	28.2386	-100.1592	3	
RAN	*Panicum hallii var. hallii*	West	USA	Texas	Lubbock	990	33.4602	-101.9143	2	
CHR	*Panicum lepidulum*		Mexico	Michoacan	Charo	1975	19.6908	-101.0312	5	
CUT	*Panicum lepidulum*		Mexico	Michoacan	Morelia	2113	19.7213	-101.3424	2	
ERO	*Panicum lepidulum*		Mexico	Michoacan	Erongaricuaro	2101	19.6072	-101.6871	5	
HTE	*Panicum lepidulum*		Mexico	Michoacan	Huaniqueo	2160	19.8959	-101.4332	3	
HUA	*Panicum lepidulum*		Mexico	Michoacan	Huaniqueo	2141	19.876	-101.4334	5	
MOR	*Panicum lepidulum*		Mexico	Michoacan	Morelia	1975	19.6761	-101.3136	5	
PTZ	*Panicum lepidulum*		Mexico	Michoacan	Patzcuaro	2128	19.5581	-101.5781	1	
TEM	*Panicum lepidulum*		Mexico	Estado de Mexico	Temamatla	2344	19.1887	-98.8921	5	
TEN	*Panicum lepidulum*		Mexico	Michoacan	Huaniqueo	2207	19.8947	-101.4284	3	
TZI	*Panicum lepidulum*		Mexico	Michoacan	Morelia	2127	19.7908	-101.4445	1	
ALD	*Panicum sp*		Mexico	Tamaulipas	Aldana	269	22.863	-98.2296	1	

Table 1.2. Morphological traits mean (standard error) for *Panicum hallii* var. *filipes* and *Panicum hallii* var. *hallii* measured in two different seasons and three localities.

Trait	**var. *filipes***	**var. *hallii***
Panicum length (cm)	24.75 (2.44)	26.38 (5.22)
Primary branch number	13.83 (4.05)	9.72 (1.5)
Secondary branch number	45.25 (28.39)	17.31 (9.10)
Biomass (gr)	247.76 (82.28)	151.92 (52.83)
Senescence score	1.81 (0.16)	2.43 (0.27)
% electrolyte leakage	15.67 (9.67)	33.89 (12.39)
Seed mass (mg)	0.77 (0.18)	1.41 (0.65)

Table 1.3 Correlation between geographical distance and genetic diversity (He, heterozygosity) in *Panicum hallii* variety *hallii* genetic clusters from the diversity hotspot.

Genetic cluster	Number of samples	Correlation	*p*-value
Central Texas	90	-0.1377	0.1954
West of the distribution	187	-0.3568	5.5E-07
Sympatry area	32	-0.2335	0.1982
Texas-Mexico	87	-0.0722	0.5060

Table 1.4 Membership probability of inferred ancestry component for individuals of *Panicum* var. *hallii* collected in the sympatric zone using four sets of *Panicum* var. *filipes* learning samples.

Individual	var. filipes learning samples set 1		var. filipes learning samples set 2		Removing var. filipes misclassified samples		All var. filipes as a learning samples	
	Fil	**Hal**	**Fil**	**Hal**	**Fil**	**Hal**	**Fil**	**Hal**
BEE_1AG	0.55	0.45	0.49	0.51	0.35	0.65	0.57	0.43
BEE_3F	0.13	0.87	0.12	0.88	0.04	0.97	0.14	0.86
GNZ_22F	0.42	0.58	0.38	0.63	0.48	0.52	0.43	0.57
GNZ_24F	0.58	0.42	0.52	0.48	0.22	0.78	0.60	0.40
RAR_8F	0.33	0.67	0.28	0.72	0.03	0.97	0.33	0.67
UVL_1AG	0.27	0.73	0.24	0.76	0.57	0.43	0.28	0.72

Table 1.5 Ploidy level measure by flow cytometry profile from some *Panicum* section *Diffusum* species collected.

Species	Genotype	Ploidy	IP	Absolute DNA
Panicum hallii* var. *filipes	CAM.1.G1	2X	33903	0.54
Panicum hallii* var. *filipes	FIL.2	2X	34733	0.55
Panicum hallii* var. *filipes	FIL.2	2X	34368	0.54
Panicum hallii* var. *filipes	FIL.2	2X	28795	0.46
Panicum hallii* var. *filipes	FIL.2	2X	27611	0.44
Panicum hallii* var. *filipes	FIL.2	2X	30029	0.48
Panicum hallii* var. *filipes	FIL.2	2X	29797	0.47
Panicum hallii* var. *filipes	FIL.2	2X	26561	0.42
Panicum hallii* var. *hallii	MCR.19.G2	2X	37359	0.59
Panicum hallii* var. *hallii	HAL.2	2X	32020	0.51
Panicum hallii* var. *hallii	HAL.2	2X	31810	0.50
Panicum hallii* var. *hallii	HAL.2	2X	31783	0.50
Panicum hallii* var. *hallii	HAL.2	2X	32231	0.51
Panicum hallii* var. *hallii	HAL.2	2X	28393	0.45
Panicum cf. lepidulum	TEN.1.G1	4X	83392	1.33
Panicum cf. caplillare	HAR.29.F	4X	93630	1.49
Panicum cf. diffusum	ENR.4.F	4X	65811	1.05
***Panicum* sp.**	ALD.1.G1	4X	87424	1.39
***Panicum* sp.**	ALD.1.G1	4X	62545	1.00
***Panicum* sp.**	ALD.1.G1	4X	60685	0.97

FIGURES

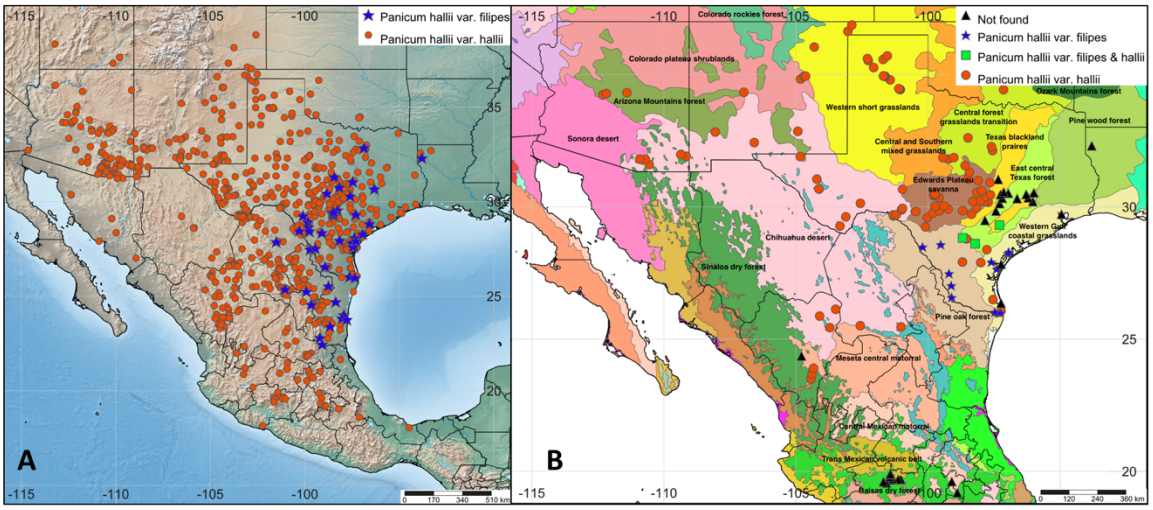

Figure 1.1: **A**. Map of the natural distribution of *Panicum hallii* varieties from GBIF occurrence download (GBIF.org (05/12/2014), herbarium visits and personal field observations. **B**. Map of collections of *Panicum hallii* and other *Panicum* species used in this study. Ecoregions showing the ecological preferences for *Panicum hallii*. The East boundary of the distribution of *Panicum hallii* matches with the East Central Texas forest, where several failed collection trips were made (black triangles). The South distribution was limited to the Meseta Central Matorral in Mexico (Olson *et al*., 2001).

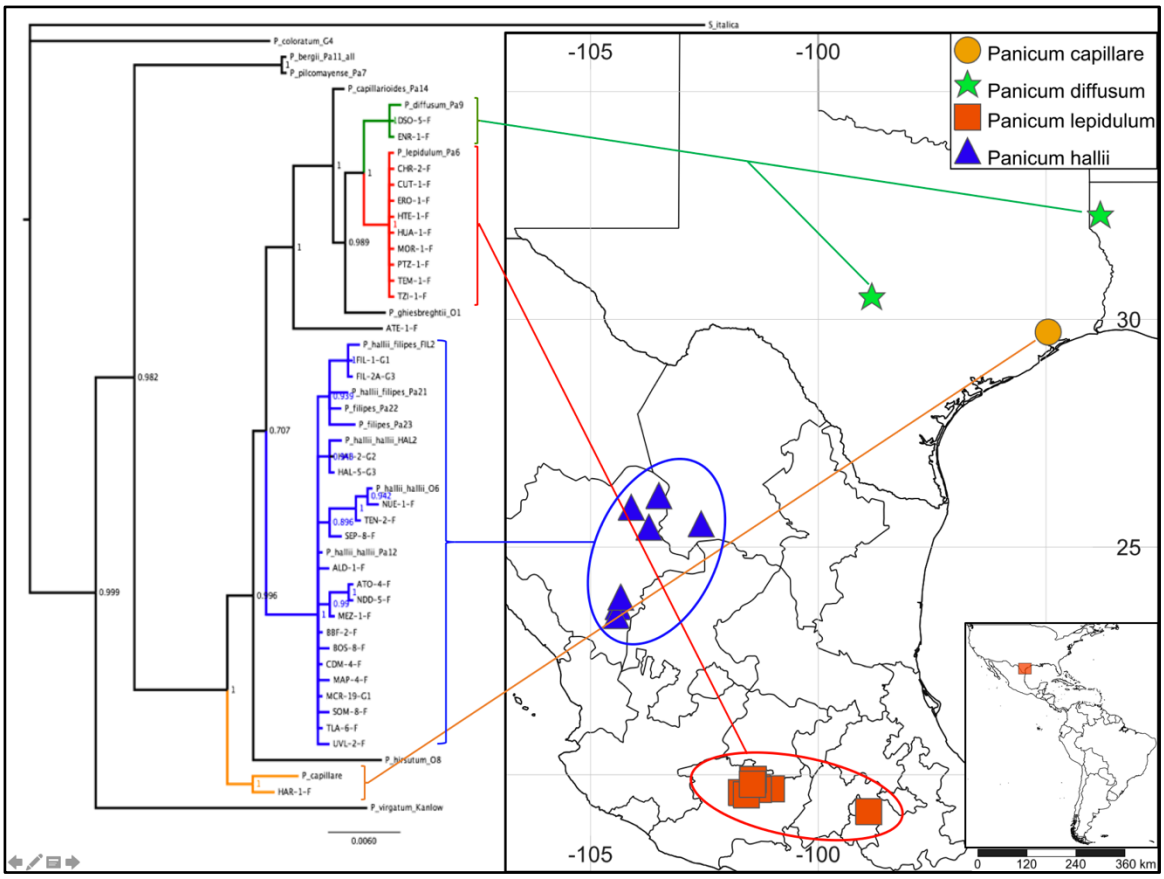

Figure 1.2. Sequences for the ITS nuclear marker and two chloroplast regions distinguish the different species from the *Panicum* section *diffusa*. This Bayesian tree based on nucleotide sequences from ITS and two chloroplast markers for samples analyzed by ddRAD (Branch support values is a consensus from 1,000 bootstrap replicates). Green lines indicate *P. diffusum* clade. Red lines indicate *P. lepidulum* samples. Blue lines indicate *P. hallii* collections. Orange clade indicates *P. capillare* collections.

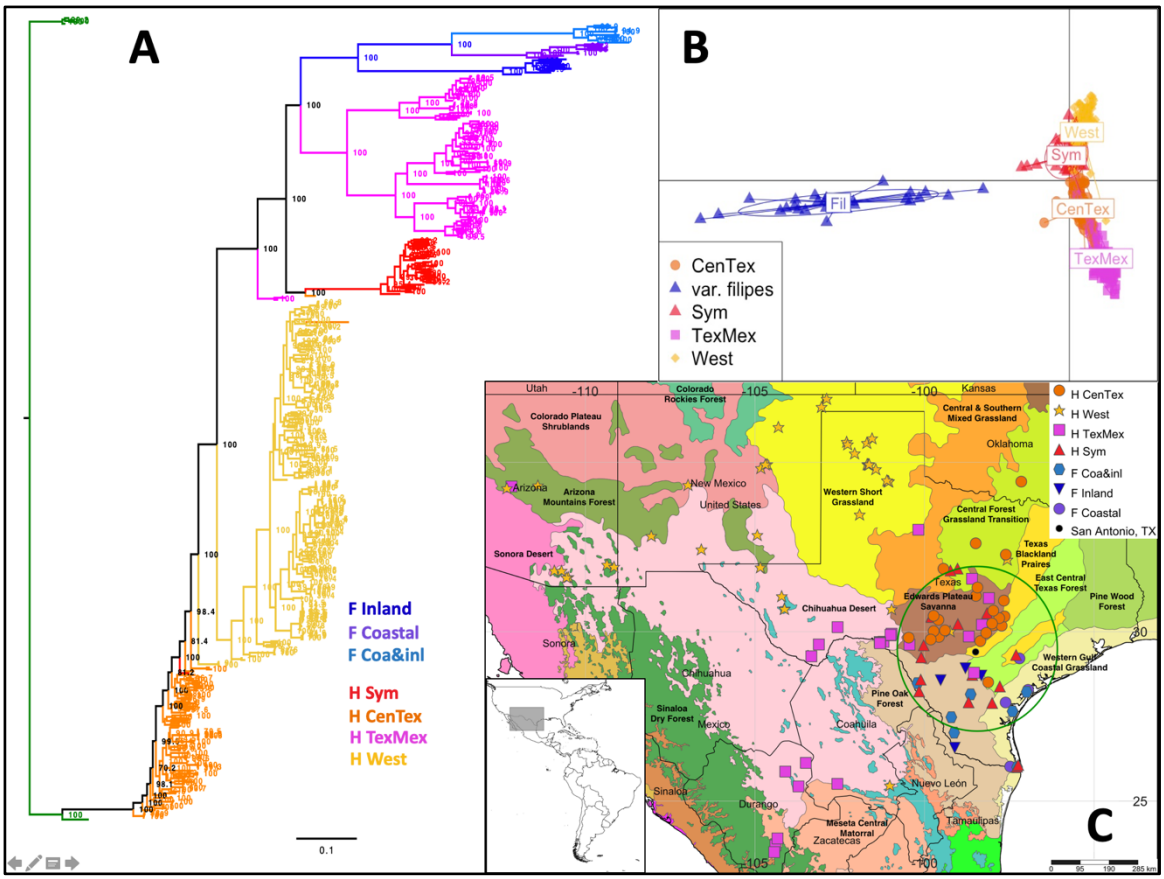

Figure 1.3. Geographic and genetic structure of 423 samples of *Panicum hallii*. A. Maximum likelihood phylogenetic tree, with the blue cluster representing *Panicum hallii* var. *filipes* and the black cluster representing *Panicum hallii* var. *hallii*. *Panicum capillare* was used as an outgroup (Branch support values after 10,000 ultrafast bootstrap replicates). B. Discriminant analysis of principal components (DAPC). C. Geographical representation of the phylogenetic tree and DAPC analysis using an ecoregion map. The green circle represents the *P. hallii* diversity hotspot 200 km around San Antonio, Texas. Maps produced using *SimpleMappr* (Shorthouse, 2018).

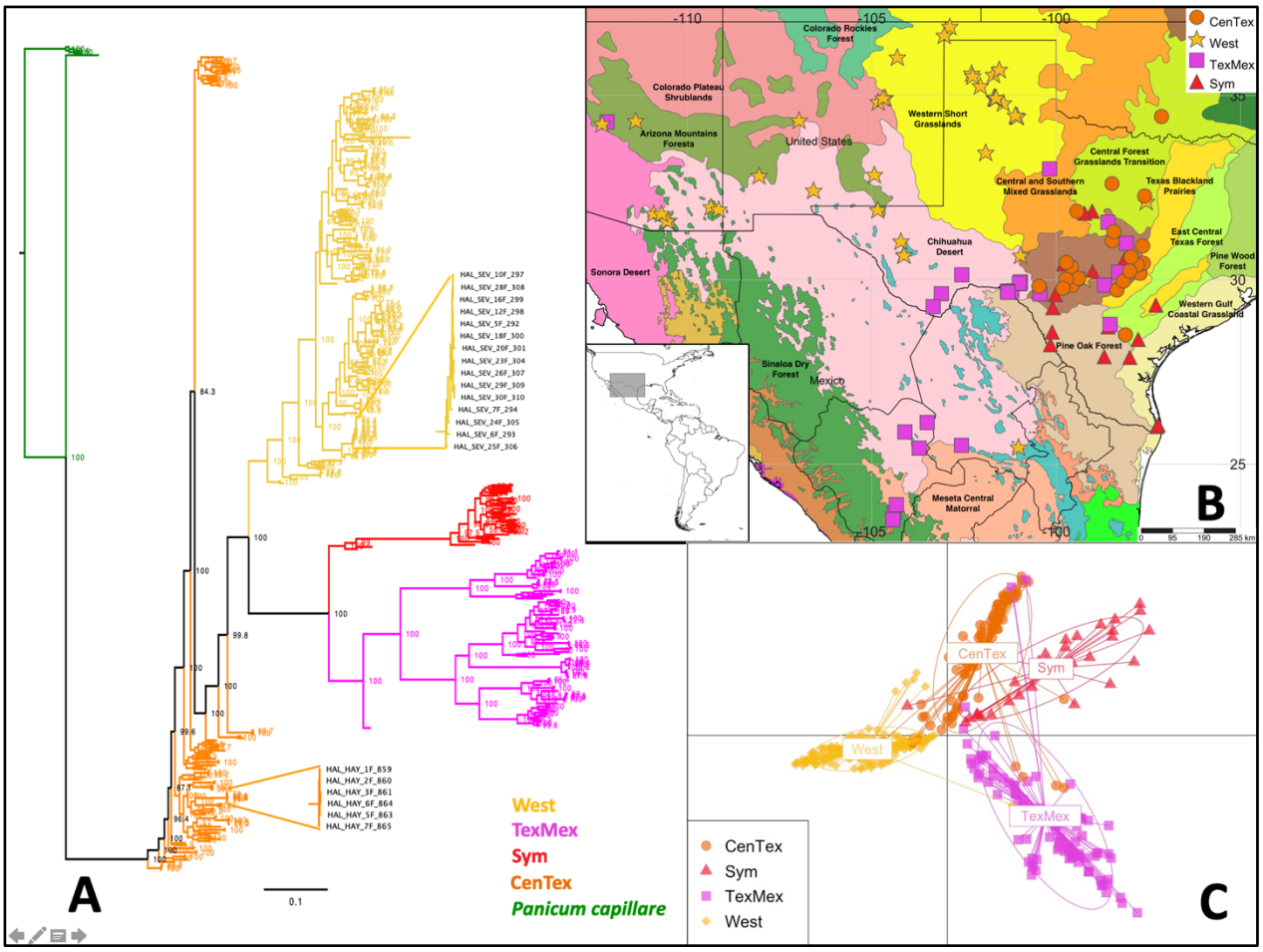

Figure 1.4. Geographic and genetic structure of 396 samples of *Panicum hallii* var. *hallii* from 104 localities. A. Maximum likelihood phylogenetic tree including inset clades depicting the close relationship between individuals from single collection locations. *Panicum capillare* was used as an outgroup (Branch support values after 10,000 ultrafast bootstrap replicates). B. Geographical and ecological representation of the clustering analysis. C. Discriminant analysis of principal components (DAPC).

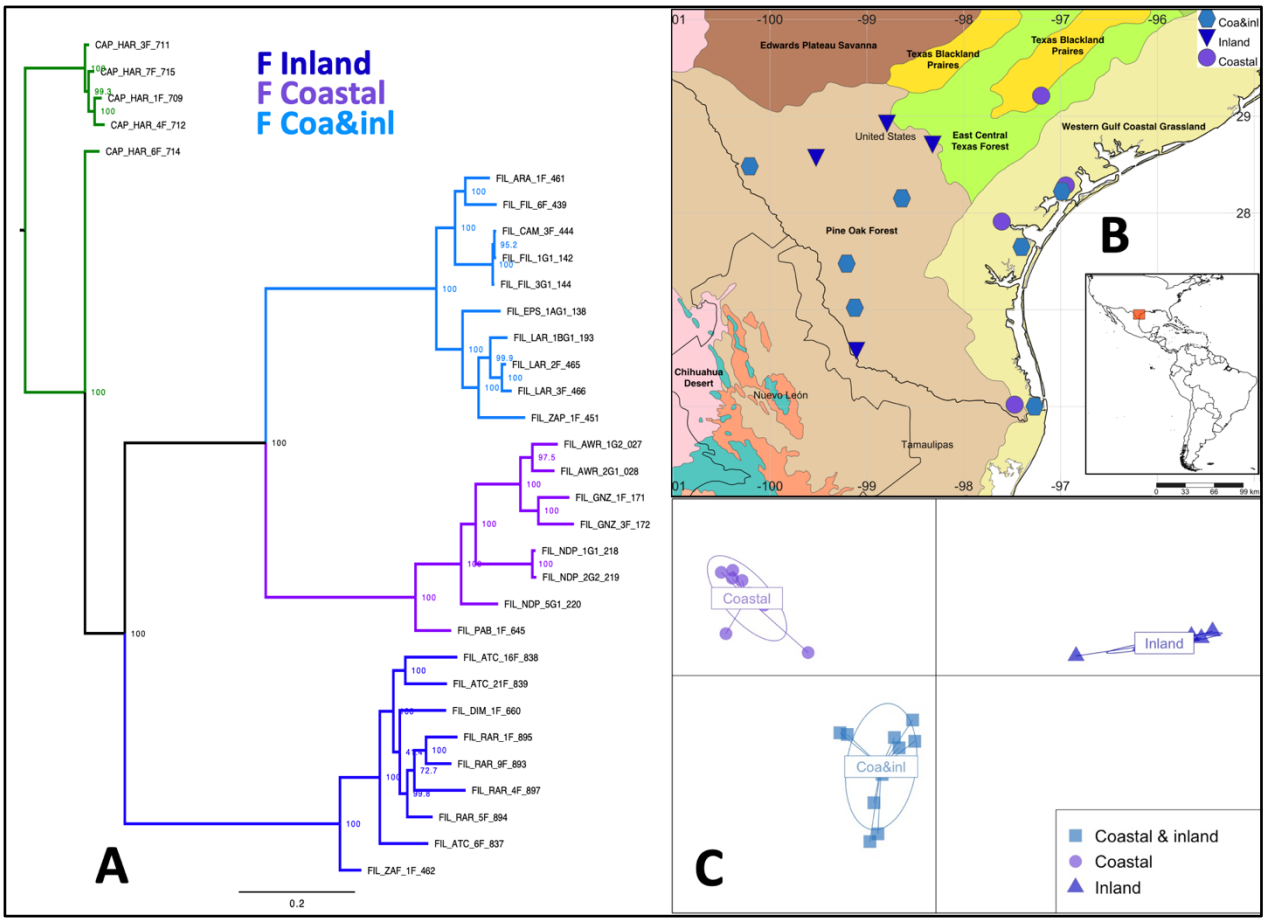

Figure 1.5. Geographic and genetic structure of 27 samples of *Panicum hallii* var. *filipes* from 14 localities. A. Maximum likelihood phylogenetic tree (Branch support values after 10,000 ultrafast bootstrap replicates). *Panicum capillare* was used as an outgroup. B. Geographical and ecological representation of the clustering analysis. C. Discriminant analysis of principal components.

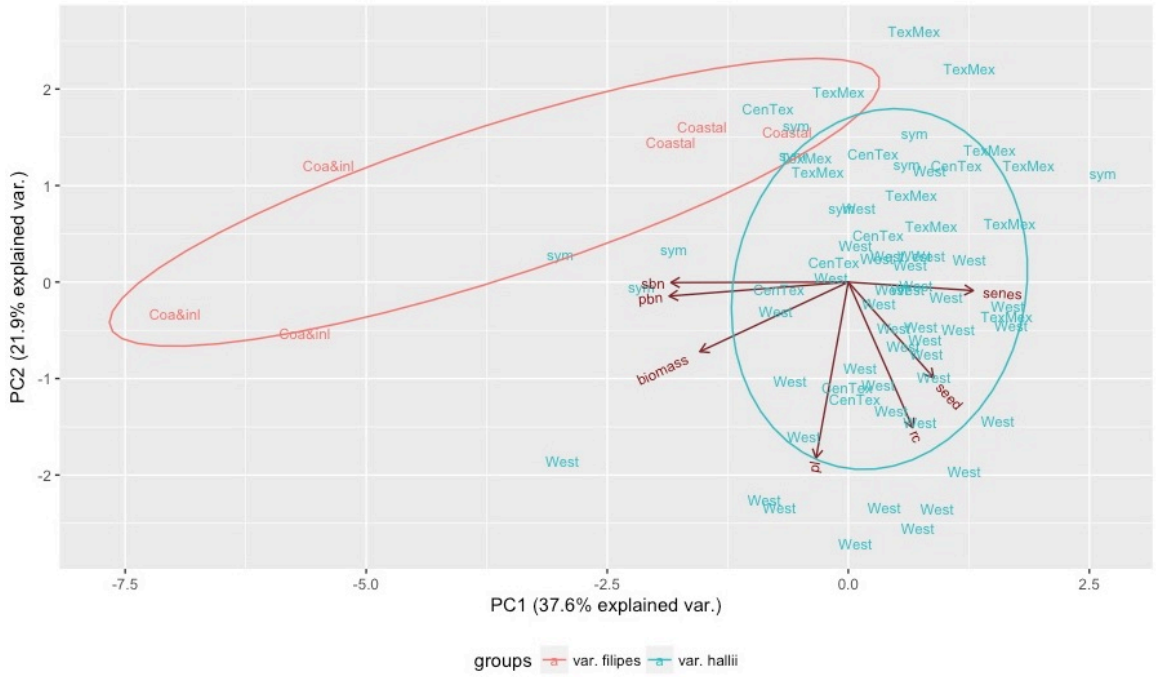

Figure 1.6 Principal component analysis of phenotypic data in *Panicum hallii* var. *filipes* and var. *hallii* (Labels represent the genetic cluster).

Chapter 2: Quantitative Trait Loci Mapping and Global Transcriptome Profiling Reveal Candidate Genes Regulating Cold Responses in the C_4 grass *Panicum hallii*

ABSTRACT

Unlike animals, plants cannot migrate from unfavorable conditions. Instead, they developed different strategies to survive by acclimating, adapting, or avoiding stress. Adaptation of plants to winter weather often involves the evolution of complex quantitative traits involving polygenic variation. In this study, we integrated complementary approaches and experimental designs to understand the response to low temperature stress in C_4 grass *Panicum hallii*. We used QTL mapping to dissect the genetic architecture of genome regions associated with cold tolerance traits and complemented this work withan analysis of global transcriptome responses to chilling. This combined approach allowed us target a subset of 36,125 candidates genes in the annotation reference genome to just 11 DEG present in a QTL interval associated with survival following freezing. A number of these genes have been associated with chilling and cold responses in other systems. Our results suggest a very dynamic transcriptome across diurnal chilling and perhaps an important interplay between light signaling and responses to decreasing temperatures.

INTRODUCTION

Given that plants are sessile organisms, they are continuously challenged by both biotic and abiotic stresses in natural environments. As such, they experience strong selection pressures to acclimate or adapt to local conditions. A major abiotic stressor affecting plant growth and development, and thereby limiting distribution of plants and crop production, is low temperature (Shi *et al*., 2018; Xu & Chong, 2018). It is estimated that crop yields across 57% of the global land area are affected by cold stress (Cramer *et*

al., 2011). The constraints resulting from cold temperatures are not a surprise, considering that flowering plants diversified in warm and wet environments at least 140 million years ago (Preston & Sandve, 2013). However, some groups of angiosperms have diversified in both the northern and southern hemispheres, and many species tolerate extreme winter conditions in sub-arctic regions (Li *et al*., 2012). In the grass family (Poaceae), for example, members of the subfamily Pooideae (temperate grasses) diversified in both hemispheres acquiring strategies to survive in low temperatures (Schubert *et al*., 2017). Included in this group are economically important cereals, such as winter wheat (*Triticum aestivum*), barley (*Hordeum vulgare*), oats (*Avena sativa*), and grasses like *Brachypodium distachyon* and perennial ryegrass (*Lolium perenne*).

Enormous gains have been made over several decades in understanding plant responses to low temperatures (Zhao *et al*., 2015; Amasino, 2018; Maleki & Ghorbanpour, 2018; Xu & Chong, 2018; Zhong *et al*., 2018). Considering the range of low temperatures plants face, cold stress in plants can be broadly classified into chilling and freezing categories. Plants subject to temperatures ranging from 0 to 20 ºC can experience chilling stress while freezing exposure occurs at temperatures below 0 ºC (Sinha *et al*., 2015; Shi *et al*., 2015). Plant responses to chilling and freezing conditions have been divided into avoidance and tolerance mechanisms. Annual life spans and the ability to prevent ice from forming are examples of avoidance mechanisms. Alternatively, plants have developed physiological and morphological characteristics to tolerate cold exposure, which can broadly be divided into two categories: vernalization and cold acclimation. Biennial and winter annual herbaceous plants from temperate regions with periods of extreme cold must synchronize their reproductive stages with favorable environmental conditions. This ability is called vernalization (Kim *et al*., 2009; Woods *et al*., 2016; Amasino, 2018; Xu & Chong, 2018). In *Arabidopsis thaliana* and temperate grasses, vernalization results in the

modification of the expression of genes that regulate flowering time (Kim *et al*., 2009). Two main genes are required to confer vernalization responses: FRI (FRIGIDA) and FLC (Flowering Locus C) (Amasino, 2018; Li *et al*., 2018; Xu & Chong, 2018). In other cases, plants induce freezing tolerance mechanisms in response to low, nonfreezing temperatures through a process known as cold acclimation (Gilmour *et al*., 1988; Thomashow, 1999; Fowler & Thomashow, 2002). Cold acclimation is a sophisticated mechanism that improves plant freezing tolerance upon exposure to nonfreezing temperatures through an array of physiological and biochemical modifications (Shi *et al*., 2018). Genetic and molecular studies have identified C-repeat binding factors (CBFs, also known as dehydration-responsive element binding or DREB), transcription factors that regulate several hundred cold-responsive (COR) genes. Some studies estimate that there are approximately 4,000 COR genes (Zhao *et al*., 2015), but just 1,200 of these genes are robustly regulated by low temperature (Zhao *et al*., 2015), indicating that the remaining COR genes are regulated by early cold-expressed CBF-independent transcription factors (Shi *et al*., 2018). In grasses it has been demonstrated that many cold-responsive patterns of regulation are not conserved between species and gains or losses of regulator genes are frequent (Preston & Sandve, 2013; Zhong *et al*., 2018).

Despite this progress, we still have much to learn. The majority of the research concerning plant cold response has been carried out using the model species *A. thaliana* and temperate grasses (Thomashow, 1999). However, the cool-season grasses are just a small part of grass diversity, with over 11,000 species, most of them from the tropics and subtropical regions, including economically important crops such as rice (*Oryza sativa*), maize (*Zea mays*), sorghum (*Sorghum bicolor*), foxtail millet (*Setaria italica*) and the biofuel crop grasses, including Miscanthus (*Miscanthus* spp.), energycane (*Saccharum* spp.), and switchgrass (*P. virgatum*). But cold adaptations in these tropical and subtropical

clades are still largely unexplored (Sandve *et al*., 2011; Rodríguez *et al*., 2014; Zhang *et al*., 2017; Peixoto & Sage, 2017), and a better understanding of responses to low temperature stress in these clades may lead to relevant agricultural improvements. In addition, traditional experimental designs often explore gene expression responses at low but non-freezing temperatures, typically at 4 ºC, which may limit the scope of inference from these experiments in tropical and subtropical plants. In particular, it is possible that many of the effects observed are better considered "stress shock" given the short timescales of stress imposition and the lack of acclimation (Des Marais *et al*., 2012). More studies are needed to identify additional transcriptional regulators, gene networks, and physiological processes related to complex responses to chilling and freezing temperatures.

Hall's panicgrass (*Panicum hallii*) is a perennial C_4 grass native to North America and is a valuable candidate species for the study of cold tolerance, due to its adaptation to diverse environments across its natural distribution (Lowry *et al*., 2013; Lovell *et al*., 2018; Gould *et al*., 2018). *P. hallii* is divided into two varieties: the widespread *P. hallii* var. *hallii* (hereafter var. *hallii*) and the more geographical restricted *P. hallii* var. *filipes* (hereafter var. *filipes*). Var. *hallii* has a natural distribution that spans from southern Colorado south into Mexico and from western Arizona to eastern Texas. Var. *filipes* is typically found along the Gulf Coast and Rio Grande Valley of Texas and Mexico (Gould, 1975; Waller, 1976). Across its natural distribution, *P. hallii* grows in a variety of latitudinal and altitudinal environments with considerable diurnal and seasonal temperature variation. Its latitudinal distribution covers a range of seven USDA hardiness zones (HZ), just in the United States. In the southern distribution the average extreme minimum temperatures ranged between -4 ºC to -1 ºC while in the northern distribution it fluctuated between -20 ºC to -23.3 ºC. Also, this species has a wide altitudinal range, being collected

from sea level along the coastal shore of the Gulf of Mexico to more than 2,000 meters above sea level in the Gila and Guadalupe Mountains in New Mexico.

Panicum hallii is also an important model system for studying C_4 perennial grasses as a diploid model for polyploidy bioenergy crops. *P. hallii* is a close relative of switchgrass (*P. virgatum*) (Zhang *et al*., 2011), an important lignocellulosic biofuel crop. Although the breeding of switchgrass has been conducted since the 1950s, improving switchgrass performance faces substantial challenges due to its ploidy (tetra- or octaploid), its high levels of heterozygosity caused by its outbred mating system, and its relatively large genome size (Casler *et al*., 2011). Importantly, both switchgrass and *P. hallii* have upland and lowland ecotypes. Lowland ecotypes are bioenergy targets due to their greater yields, but these ecotypes are severely affected by the harsh winters of the northern latitudes (Casler & Vogel, 2004; Casler *et al*., 2007; Lee *et al*., 2018). As a result, breeding for winter hardiness in lowland ecotypes may increase the biomass achieved for this crop (Casler & Vogel, 2014).

QTL mapping studies have been conducted to identify genetic loci associated with tolerance to low temperatures in some cereal crops. In rice, which is particularly susceptible to cold (Sperotto *et al*., 2018), QTLs affecting the response to low temperatures were found and used in plant breeding programs to develop new high-yield cold-tolerant varieties (Li & Mao, 2018). Although several studies have identified QTL related to cold tolerance in maize, so far those studies have not resulted in the development of improved commercial varieties (Auton *et al*., 2015). This is due in part to the challenge of positively identifying functional genes after QTL mapping, given large QTL intervals that often contain many hundreds of genes (Pandit *et al*., 2010). Recently, advances in transcriptome studies with massively parallel cDNA sequencing (RNA-seq technology) have facilitated the identification of functional genes involved in pathways related to target traits (Ozsolak &

Milos, 2011). Integrating QTL mapping with transcriptome analysis has opened the possibility of combining both approaches to further refine mapping resolution and identify functional genes involved in the response to abiotic stresses.

In this study, two approaches were used to evaluate the cold-response genes of *P. hallii*. First, a QTL analysis was utilized to identify genomic regions underlying survivorship and regrowth following freezing conditions. Second, we conducted a time series experiment with a gradual drop of temperature to characterize the transcriptional response of *P. hallii* to chilling temperatures. Our experiment focused on declining temperatures and finished at the point where most of the cold research literature experimental designs start; thus, our study fills a knowledge gap regarding the early regulatory responses to chilling in plants. These approaches allowed us to characterize QTL responses to freezing conditions, identify putative sequence differences between accessions, and explore patterns of genetic variation in transcript responses to initial and mild chilling stress conditions.

MATERIALS AND METHODS

Quantitative Trait Loci Mapping

Establishment of the recombinant inbred lines mapping population

A population of recombinant inbred lines (RIL) was established to evaluate the genetic basis of the cold-tolerance responses in *P. hallii*. Genotyping and genetic map development of the RIL population were performed as described by Khasanova *et al*., 2018. Briefly, the parents of the mapping population were a sample of var. *filipes* (FIL2) which consistently showed cold tolerance measured with electrolyte leakage assay (described below) and var. *hallii* (HAL2), which was sensitive to cold stress (Figure 2.1).

A hybrid F_1 cross of these two individuals, using HAL2 as a maternal parent, was completed to create a large F_2 population through self-fertilized seeds obtained from the hybrid individual (Lowry *et al*., 2015). Lines were developed following the single-seed descent method from the F_2 to F_6 generation. Whole genomes of 356 RILs lines were resequenced, and 722 bin-markers were used to build a high-quality genetic (Khasanova *et al*., 2018).

Freezing treatment and phenotyping

In September 2015, four biological replicates of 183 RILs and 33 replicates of each parent (FIL2 and HAL2) (798 plants in total), were grown in a 3.5-inch square pot filled with six parts of Pro Mix BM (Premier Tech Horticulture, Rivieredy-Loup, Quebec) and two parts of Turface (Profile, Buffalo Grove, IL) in The University of Texas at Austin greenhouse. Greenhouse conditions were controlled with supplementary lighting to maintain a 16-hours-long day. Temperatures were controlled with evaporative coolers and fans allowing midday peak temperatures of 33 °C and 22 °C nights. To establish the freezing temperature to be tested in the study, five samples for each parent (HAL2 and FIL2) were exposed to low temperatures at -2 ºC, -3 ºC, -4 ºC, -5 ºC, and -6 ºC for 12 hours using an environmental chamber (EJS limited). These plants were then shifted to greenhouse conditions to allow for recovery and then compared to two controls held at 26 ºC under both greenhouse and growth-chamber conditions. After a month of recovery under growth-chamber conditions at 26 ºC and 12D:12L photoperiod, the survivorship of the samples was evaluated (0 = dead, no green tissue or regrowth; 1 = alive, visible regrowth and green leaf tissue). Resulting from this initial analysis, a temperature of -4 ºC was selected as a target experimental condition for our QTL studies as it showed the largest difference between HAL2 and FIL2.

On December 9 of 2015, the RILs and parent samples were transferred from 26 ºC and light conditions to a freezing treatment of -4ºC overnight (12 hours of darkness) with the four replicates divided into blocks placed in an environmental chamber. After the treatment the plants were transferred to greenhouse conditions (16-hours day and temperatures of 33°C day/22°C nights) to allow for recovery. We recorded phenotypic data at one and three months following the freezing treatment. The phenotypic data consisted of survivorship (0 = dead, no green tissue or regrowth; 1 = alive, visible regrowth and green leaf tissue); for plants that were alive, tiller number and plant height were additionally recorded as a metric of the quality of the recovered plant.

QTL mapping

For QTL mapping, the means of four biological replicates of each RIL were used as traits for six measured phenotypes in this experiment including Survivorship (S) and tiller number (TN) at one and three months after freezing treatment. QTL analysis was performed in R using R/qtl2 v. 16 package (Broman et al., 2018). First, the genotype probabilities were calculated for a 1 cM grid with the observed marker data using an error probability of 0.001. R/qtl2 fits a linear mixed model to test for marker-trait association against a null hypothesis that there is no QTL and it allows kinship to be included as random variable to control for background polygenic variation. Genome scans were performed using Haley-Knott regression (Haley & Knott, 1992) with the LOCO model (Leave One Chromosome Out), which derives the kinship matrix for each chromosome from all chromosomes except the one being tested. Including relatedness in a LOCO model minimizes over-estimation of QTL effect size and controls for background residual genetic variation. The null hypothesis distribution for each phenotype was obtained using 1000 random permutations of genotypes with respect to phenotype, which provides a genome-

wide critical value for type 1 errors. Significant LOD peaks were identified using this critical value with an alpha = 0.05 for all traits. To obtain QTL intervals, the lodint function was used to obtain a 1.5 LOD support interval around each QTL. Trait heritability was calculated with one-way ANOVA among clonal lines with 1000 bootstrap run using h2boot (Phillips & Arnold, 1999).

Transcriptome analysis to cold responses

Plant material and growth conditions

Four RILs and the two parents (FIL2 and HAL2) were selected for extensive global transcriptome analysis in response to chilling. The four RIL lines were chosen using the response to the QTL essay of freezing treatment, including two lines which did not survive and two lines in which all the four of the replicates survived were chosen. RIL line 298 and line 247 were selected as tolerant and line 113 and lines 24 as sensitive. The selection of RIL based on this criterion was driven by an interest in sampling varied responses to chilling and freezing tolerance for our transcriptome studies.

On June 23 of 2016, 40 seeds of each genotype were germinated, following the conditions described by Lowry *et. al*. (2013). On July 6 and 7 of 2016, 13 days after sowing (DAS), around 30 plants of each line were potted and transferred to a growth chamber. Plants were thinned after 30 DAS (August 6 2016) to one per pot and were grown under a short-day photoperiod (12 hours light/dark) to synchronize the flowering time between var. *filipes* (long day) and var. *hallii* (short day)(Lowry *et al*., 2015), as well as to simulate short winter days. During this establishment period, the temperature was set at 26 ºC during the day and at 22 ºC during the night. These growth temperature conditions were selected as they approximate the temperatures occurring in the growing season of *P. hallii* in Central Texas. In contrast, the 4 ºC low temperature was chosen to represent the average

temperature during the early winter season and because it is above the freezing but often acclimates plants without causing tissue damage (Zhong *et al*., 2018). Experimental plants were watered every two days so that the soil was saturated to field capacity. Plants were grown in three replicates with treatments and genotypes randomly located for a total of 144 plants.

We chose mature plants (53 DAS) for our experiment, considering that in natural populations plants at this age experience the first low temperatures early in their first winter season. Our cold treatment was applied gradually by decreasing the temperature at a rate of 1 ºC per hour, starting at 26 ºC and decreasing to 4º C in 24 hours. Over the course of our 24-hour experiment, two environmental cues were manipulated under growth chamber conditions: temperature and light. As temperature decreased through the experiment, the light conditions followed 12 hours of light and 12 hours of dark to complete a 24-hour period of circadian cycle. It is very well known that metabolic and physiological changes happen during this circadian cycle (McClung, 2006, 2015). For clarity, we report results following the zeitgeber time (ZT) method (McClung, 2006), setting ZT0 as the time when experimental lighting is initiated as representing dawn, ZT12 when the lights turn off representing dusk, and finally ZT24, when the period is complete with the transition to the following morning period.

Leaf tissue from the basal part of the flag leaf was collected at each treatment, beginning at ZT2 hours and 26 ºC, ZT8 and 21 ºC, ZT14 and 15 ºC, ZT20 and 9 ºC, and ZT2 and 4 ºC. In another growth chamber, another set of plants was grown as a control at 26 ºC using the same photoperiod (12:12 hours, light:dark throughout) and the tissue was collected as the same time points (Figure 2.2). For the treatment at Z2 and 4 ºC, the Z2 and 26 ºC was used as a control, although there were 24 hours of difference between sampling.

Under dark conditions, the tissue was collected using a low intensity red light to minimize light exposure effects.

Phenotypic response to cold treatment

To evaluate the phenotypic response to cold stress, electrolyte leakage assays were conducted before (48 DAS) and after (61 DAS) the experiment. Under freezing conditions, the plasma membranes in plant tissue are disrupted and release ions to the extra cellular medium (Steponkus, 1984). Electrolyte leakage can be quantified by soaking temperature-treated leaves in water and measuring the electrical conductivity of the water using a standard lab conductivity meter. This experimental protocol is commonly used to quantify freezing injury to plant tissue (Campitelli *et al*., 2013; Pérez-Harguindeguy *et al*., 2013). We collected 4 pieces of 4 cm each from a basal flag leaf in a 15 ml centrifuge tube for each plant. These samples were rinsed with deionized water (diH_2O) and surface dried. We then exposed the tissue to -4 ºC for 4 hours using a temperature-controlled chamber (Thermotron 3200, Holland, Michigan, USA). To allow the tissue to adjust to freezing conditions, the temperature was gradually reduced from room temperature (24 ºC) to the target temperature (-4 ºC) over the course of one hour. Tissue was gradually thawed after treatment over the same time period. We induced the formation of ice crystals via ice nucleation. Following the temperature treatment, 8 mL of diH_2O were added to each tube, and the tubes were placed in a shaker at room temperature for 24 hours to allow the leakage ions to equilibrate with the diH_2O. Then, the treatment conductivity (C_{TREAT}) was taken, using a conductivity meter (VWR SympHony Model H10C). Next, the tissue cells were lysed in a boiling water bath for 30 minutes, and subsequently allowed 24 hours in the shaker to release all remaining ions. Finally, the second conductivity measurement was taken (C_{TOTAl}). Relative conductivity (RC) was calculated as a proportion of the C_{TREAT} with respect to the

C_{TOTAL} (RC = C_{TREAT} / C_{TOTAL}). RC ranged from zero (no damage; freezing-tolerant) to one (complete damage; freezing-sensitive). A linear model was fit (relative conductivity = μ + treatment + genotype + treatment-by-genotype) to test for genetic and treatment effects on electrolyte leakage using JMP® Genomics v. 8.0.

Leaf tissue collection and RNA sequencing

Tissue for RNA extraction was collected from the three independent biological replicates of each treatment and control plant as follows: (1) Three tillers that were representative of the canopy were selected from each plant, (2) The fully expanded flag leaf was removed at the ligule level for each tiller, and (3) ~2 cm from the ligule for the three samples were cut into small pieces and combined in a single 2 mL microcentrifuge tube with three stainless steel beads on it. Immediately after collection, the tissue was frozen in liquid nitrogen and preserved at -80 ºC until RNA extraction was conducted. The scissors, beads, and gloves used in sampling tissue for the experiment were treated with RNAaseZap® (Ambion) to remove RNase contamination. Before the RNA extractions, the tissue was ground to fine powder, using a Geno/Grinder 2000 for 30 seconds at 700 rpm. Total RNA was extracted following the standard TRIzol® (Ambion) protocol followed by DNase I treatment to remove contaminating genomic DNA. RNA quality and purity were evaluated in 1% agarose gel and the quantity was assessed using a Qubit™ (Invitrogen) fluorometer and the RNA high sensitive assay.

One μg of intact total RNA per sample was used to prepare cDNA tag libraries using a modified version of 3'-TaqSeq protocol developed in corals (Meyer *et al*., 2011) and applied to *Panicum* species (Meyer *et al*., 2012, 2014; Lovell *et al*., 2016). The reliability of 3'-TagSeq has been evaluated by qPCR (Meyer *et al*., 2012) and also compared with whole mRNAseq method (NEBNext®)(Lohman *et al*., 2016). Taken

together, these evaluations show that the 3'TagSeq protocol is an efficient low-cost alternative to evaluate relative gene expression in a large number of samples. In short, purified 3' RNA was amplified and sample-specific oligonucleotide barcodes were assigned to the Illumina HiSeq platform adaptors to pool the samples. Prepared libraries were submitted to the Genomic Services Lab at HudsonAlpha Institute for Biotechnology and sequenced on the Illumina Hiseq-2500 platform. We aimed to obtain sequencing coverage of five million 150 bp single-end reads per sample.

Bioinformatic Analysis of RNA-Seq Data

The quality of the raw individual nucleotide sequences (reads) for each sample in was evaluated using FastQC v. 0.11.15 (Andrews, 2018). Raw reads were trimmed to remove sequence adapters, homopolymers (poly-A or poly-T with a minimum length of 20 bps) using Cutadapt v. 1.14 (Martin, 2011). Reads were retained using a minimum average sequence quality of 20 and a minimum length of 70 bps. An analysis of the k-mer profile of *P. hallii* 3' prime transcriptome revealed that 90% of *P. hallii* transcripts can be uniquely detected with a read length of 100 bps (T. Haque, personal communication). These filtered reads were subsequently aligned to the *P. hallii* var. *filipes* v. 2 reference genome (DOE-JGI, https://phytozome.jgi.doe.gov), using BWA MEM algorithm (v. 0.7.16a-r1181) (Li & Durbin, 2010). In order to remove ambiguous alignments, reads were filtered with a minimum mapping quality score of 10 using SAMtools (v. 1.5) (Li *et al.*, 2009). Finally, transcript counts were generated from the filtered alignments using FeatureCounts (Liao *et al.*, 2014) included in Subread package v. 1.6.2 (Liao *et al.*, 2013) and utilizing the gene annotation corresponding to *P. hallii* var. *filipes* v. 2 reference genome (DOE-JGI, https://phytozome.jgi.doe.gov).

Statistical Analysis of RNA-Seq Data

The statistical analysis began with a pre-filtering step to remove genes with low counts. The raw count matrix was culled to genes with mean counts of < 1. Then, in order to explore the quality of the gene count data and test for potential outlier samples, principal component analysis (PCA) of a variance stabilized count data was used to visualize the gene expression of samples and clustered by each treatment (Temperature/ZT) versus control. The Bioconductor R package DESeq2 (Love *et al*., 2014) was then used to test for differential gene expression, using a negative binomial distribution for gene counts and a likelihood ratio test (LRT) to evaluate statistical significance for experimental factors including genotype, temperature treatment, and their interaction at each studied zeitgeber time. For instance, the interaction effect between genotype and treatment was tested using linear models: a full model including treatment, genotype and treatment by genotype interaction was compared to a reduced linear model including only treatment and genotype. The direction of fold change was set using the 26 ºC control as a reference for treatment and treatment-by-genotype effect and the HAL2 genotype was used as reference while setting contrasts for genotype effects. We used a false discovery rate of 0.05 with the Benjamini-Hochberg approach to correct for multiple-hypothesis testing (Love *et al*., 2014).

In all cases, these analyses were conducted split by ZT time. Volcano plots were made to characterize the differential expression between treatment and control in each time/temperature point where the log2 fold change of treatments compared to control was plotted on the horizontal axis and the *P*-value of the associated test was on the vertical axis. To compare different gene sets, the R package VennDiagram (Chen & Boutros, 2011) was used to generate venn diagrams. We were especially interested in combining information from our QTL studies with the data from our transcriptome studies. The physical positions

of markers flanking each QTL were used to obtain a list of annotated genes occurring in each QTL interval. A customized R script was used to find the overlapping genes that showed significant differential expression for the interaction of genotypes and treatment and were physically annotated in a given QTL interval.

Gene Ontology Enrichment Analysis

Gene ontologies (GO) for differentially expressed genes (DEG) were retrieved from the annotation for the genome to *P. hallii* var. *filipes* v. 2 from Phytozome v.12.1.6 (DOE-JGI,https://phytozome.jgi.doe.gov). In our study, while testing for enrichment for GO terms, we used the background annotation for only those genes that were expressed in the current experiment from leaf tissue. The R package TopGO (Alexa & Ranhnenfuhrer, 2016) was used to conduct the gene ontology enrichment test using a Fisher's exact test (*p*-values < 0.05) for the three GO domains: Biological processes (BP), molecular functions (MF), and cellular component (CC).

RESULTS

Phenotype and QTL mapping for freezing response

As expected from the parental phenotypic response to freezing tolerance, we observed large variation in survival to freezing conditions in the RIL population (Table 2.1). For example, 45% (SE±0.08) of the FIL2 plants survived the -4 ºC treatment while only 21% (SE±0.08) of the HAL2 plants survived (*p value* = 0.031). Overall, 49% of the RIL survived. We detected a single common QTL on chromosome 9 with effects on survivorship and tillering (Figure 2.3, Table 2.2). As expected, the FIL2 alleles (AA) improved survival after one month and tiller number after one and three months following

the freezing treatment (Figure 2.4). The percent variance explained by the chromosome 9 QTL ranged from 7.26 to 8.52 for survivorship and tillering respectively.

Electrolyte leakage of *Panicum hallii* in response to cold temperatures

An electrolyte leakage test was completed before and after cold treatment at 4ºC as a proxy of physiological responses to chilling. We detected a main effect of the temperature treatment on electrolyte leakage, suggesting a general acclimation response (*p-value* < 0.0001). The non-acclimated plants had significantly more electrolyte leakage (Mean = 62.46, SE±2.08) than acclimated plants (Mean = 44.72, SE±2.1) (Figure 2.5). The relative conductivity before the cold treatments presented a variable distribution of responses, with high frequencies of samples suffering excessive damage (Figure 2.5A). In contrast, after cold treatments following an acclimation treatment, the relative conductivity was more homogeneous across the distribution, showing that the cold hardening increased the freezing tolerance of the individuals (Figure 2.5B). A marginally significant effect was found for genotype but we did not detect significant effects of genotype-by-treatment interactions on electrolyte leakage (*P-values* 0.0679 and 0.1362 respectively).

Transcriptome sequencing and gene counts

To investigate transcriptome dynamics during a gradual decrease of temperature, we performed Tag-Seq experiments using total RNA isolated from leaf tissue. More than 1.5 billion raw reads were obtained from the 144 samples under study. On average, each sample received 6,623,600 reads (StDv 2,761,416), of which 83% of which passed the stringent quality filtering (1.2 billion high-quality reads in total, on average 5,487,682 per sample, SD 2,336,305). After mapping the high-quality reads to the reference genome,

20,019 unique loci (55.41%) from the 36,125 loci present in the annotation file (*P. hallii* var. *filipes* v. 2 reference genome) were recovered.

Global comparison of transcriptome reveals a relationship among cold treatments and genotypes

To explore the multivariate pattern of expression variation we examined PCA plots split by each ZT/temperature combination (Figure 2.6). Genotype has a striking impact on expression across all ZT/temperature treatments with clear separation of parental genotypes. As perhaps expected, RIL lines hold an intermediate position in PC space relative to the parental genotypes. The initial stages of the experiment (21 ºC) reveal no strong differentiation between the treatment and the control (Figure 2.6A). At ZT14 and15 ºC, more significant expression differences emerge between treatment and control (Figure 6B) and continued to be more evident in subsequent conditions (Figure 2.6C). Finally, at ZT2, when the temperature reaches the coldest condition (4 ºC), the differences between treatment and control were strong.

Differentially Expressed Genes by Treatment Effect

At the false discovery rate (FDR) of 5%, we discovered 13,514 unique differentially expressed genes across four different target temperatures when compared to 26 ºC controls. Early in the chilling response, after 6 hours of treatment (ZT8) and with only a 5 ºC decrease, 1,208 genes were differentially expressed, with a majority of responses down regulated (744 genes, 61.6%). After 12 hours (ZT14) and at 15 ºC, the differentially expression genes increased to 8,598 with a relatively even balance between up-regulated (4,268 genes, 49.63%) and down-regulated responses (4,330 genes, 50.37%)(Figure 2.7). At 9 ºC, after 18 hours (ZT20) we detected 1,751 DEG with a majority of them being down regulated (1,064, 60.8%). Finally, at the lowest temperature level in the study, 4 ºC, and

after a complete 24 hours of treatment (ZT2), closing the circadian cycle, the DEG increased to 8,859 genes with an almost equally down and up regulation of responses (4,611 genes, 52% and 4,248, 48%, respectively) (Figure 2.7, table 2.3). Considering the dynamic gene expression responses across the breadth of our test temperatures, on average 30% of the DEG were exclusive or unique to each temperature/ZT treatment and just 69 genes (0.5%) were shared by all four treatments. Considering only those genes that were consistently down or up-regulated, just two down-regulated genes were shared by all four treatments, and none of the up-regulated genes were shared between treatments.

Gene ontology (GO) enrichment analysis of DEG at the three levels of functional classification (MF, molecular function; CC, cellular component; and BP, biological process) identified 237 GO terms which showed significant enrichment across the four temperature/ZT treatments. Although GO terms associated with oxidative stress start to be present at 21ºC/ZT8 and plasma membrane organization at 15 ºC/ZT14, it is only at 4 ºC/ZT2 that the GO terms directly associated with abiotic stimuli are enriched.

Differentially Expressed Genes at Genotype-by-Treatment Interaction

An underlying hypothesis is that freezing-tolerance differences between *hallii* and *filipes* may result from differential acclimation related to differential gene expression in response to chilling. As such, we were especially interested in asking the degree to which expression differences were constitutively different among genotypes, the degree to which decreasing temperature induced or repressed expression, and whether genotypes differed in their expression responses to chilling. At an FDR of 5%, we detected 581 genes that exhibited significant genotype-by-treatment interactions across the sampled ZT periods. Most of these genes were exclusive for each temperature/ZT treatment (92.8%). In the analysis of GO enrichment, in total 116 GO terms were found, some of them associated

with stress responses, especially at 4 ºC/ZT2 treatment in the context of genotype-by-treatment interactions (Figure 2.8).

Differentially Expressed Genes Located in QTL Intervals

A major goal of our work was the identification of the genes underlying freezing tolerance among var. *hallii* and var. *filipes*. A major result of our study comes from the genes localized to QTL intervals, especially those differentially expressed or that exhibit expression polymorphism among genotypes and in response to low temperatures. This approach allowed us to search for candidate genes involved in cold tolerance across the time series treatments. In the overlapping region of chromosome nine a total of 1,710 genes were localized and 944 genes were found in the QTL interval for plants which survived after a month of the freezing treatment (Table 2.4). GO enrichment analysis for the 944 genes localized in the smallest overlapping QTL showed enrichment for 24 GO terms including "response to abiotic stress".

This list of genes was further reduced by considering only genes exhibiting significant genotype-by-treatment and occurring within the QTL interval. 28 unique genes were identified based on this criteria and show a diversity of differential responses to the chilling treatment. We did a careful literature review relating these genes to chilling and low temperature responses. Surprisingly, 22 of these genes were associated with temperature responses in at least one published study (Table 2.5). Among this small list of 28 genes, 11 GO terms were significantly enriched including photosynthesis and light reaction biological processes.

Interestingly, these 28 candidate genes were largely differentially expressed in unique temperature/ZT treatment combinations – a pattern observed across most contrasts in our data. Five genes were differentially expressed at the mildest treatment of 21 ºC.

Pahal.I01704 is an interesting candidate that is directly associated with vernalization, specifically with the K-box region and MADS-box transcription factor, a potent flowering repressor (Amasino, 2018; Xu & Chong, 2018). At 15 ºC, we detected a candidate gene associated with membrane stability. At 9º C we detected three genes that were differentially expressed – these genes are of special interest because they are novel and do not have annotation and therefore may represent novel genes of unknown function. Finally, as was expected, most of the genes (21 or 67.8%), were found at the lowest temperature treatment (4 ºC). The majority of these genes have functional annotations or literature support for functions associated with cold responses in plants (Table 2.5 and Figure 2.9).

DISCUSSION

In this study, we integrated different approaches and experimental designs to understand the response to low temperature stress in grass the C4 perennial *P. hallii*. Here, we combined the power of QTL mapping to dissect the genetic architecture of genome regions associated with cold tolerance traits with an analysis of the global transcriptome to understand the mechanistic cold responses in *P. hallii*. This combined approach allowed a winnowing from 36,125 genes in the annotation reference genome to just 11 DEG present in a QTL interval associated with survival following freezing. A number of these genes have been associated with chilling and cold responses in other systems. Our results suggest a dynamic transcriptome across diurnal chilling and perhaps an important interplay between light signaling and responses to decreasing temperatures.

Co-localization of two QTL intervals associated with freezing tolerance

The plant capacity to tolerate low temperatures is a quantitative trait (Humphreys & Gasior, 2013) and is usually consider polygenic. However, our QTL mapping analysis

revealed a simple genetic architecture with a single genomic region on chromosome 9 impacting survivorship and regrowth after freezing stress. As such, we anticipate important genes involved in freezing tolerance must occur with this interval.

Transcriptome changes across gradual temperature drop experiment

Traditional transcriptome studies expose the plants directly to temperatures close to the freezing point, usually at 4 ºC, and the tissue sampling is done in a time series (Fowler & Thomashow, 2002; He *et al*., 2015; Zhong *et al*., 2018). However, because *P. hallii* is a subtropical plant, naturally distributed in some areas where the winter season is mild, we considered an experimental design where the temperature was gradually dropped across time to elucidate more realistic transcriptome responses during low temperature conditions. Following this approach, each temperature/ZT treatment had a relatively distinct set of transcripts. For example, models evaluated treatment and treatment-by-genotype effect consistently revealed unique DEG and GO analysis for each treatment (Figures 2.7, 2.8, and 2.9). Second, all the genes differentially expressed in the QTL intervals were exclusive for each ZT/treatment combination. Together, these pieces of evidence might suggest that the cold temperature regulatory network in plants is a complex process, which can be triggered even under early cool conditions. However, it is also necessary to consider the connection of circadian clock regulation, which is widely known to have an effect on cold responses and flowering time regulation (Zhong *et al*., 2018).

Early chilling stress triggers the cold acclimation transcriptome in *Panicum hallii*

Although the first temperature/ZT treatment in our experimental design was at 21 ºC, just 5 ºC degrees below the normal growing conditions, an early response eliciting cold acclimation genes was observed. In the complex gene expression network to response to

low temperature stress, the plasma membrane appears to be the perception center for temperature changes (Sinha *et al*., 2015), changing its configuration to prevent cell disruption during cold stress by stabilizing the membrane lipids. As a consequence, in our experiment, under cool conditions, 154 and 129 genes were involved in cellular lipid metabolic and lipid biosynthetic process terms, respectively. Genes involved in lipid metabolism were consistently regulated in plants of maize and sorghum, species belonging to sister tribe Andropogoneae, under cold stress at 10 ºC (Zhang *et al*., 2017). Another early response to cold, the response to oxidative stress due to generation of reactive oxygen species (ROS) was associated with 798 annotated genes belonging to an oxidation-reduction process. In addition, one gene associated with vernalization in *A. thaliana*, rice, and grasses (AGL7, AP1, MADS-box family gene with MIKCc type-box)(Woods & Amasino, 2015; Woods *et al*., 2016; Xu & Chong, 2018) present in the QTL interval was upregulated. Although vernalization-mediated flowering regulation is an important reproductive strategy in temperate grasses, evolutionary evidence suggests that rudimentary cold response networks evolved early in flowering plants (McKeown *et al*., 2016; Schubert *et al*., 2017; Zhong *et al*., 2018). These gene expression patterns suggest that even a small decrease in temperature might trigger a set of transcriptomic processes to prepare the plants for cold or even freezing temperatures. For example, irreversible damage to cold injury has been reported in bananas exposed to 12 ºC (Yang *et al*., 2015), so is presumably those elicit gene involved in cold acclimation can be expressed at warmer temperatures.

Surprisingly, at 15 ºC, only genes involved in ROS activity were found in the GO enrichment analysis with 948 genes associated with oxidoreductase activity terms. In addition, none of genes associated with this temperature/ZT from the QTL interval were

annotated. This might be an indication that these genes can be important in the cold adjustment in monocot plants.

Water uptake and membrane transport capacity are one of the processes enhanced during cold responses in non-temperate grasses (Dametto *et al*., 2015). GO enrichment analysis of the DEG in the genotype-by-treatment model at 9 ºC reveled six molecular function terms, all of them associated with transporter activity, especially at the transmembrane level. Transporter capacity during cold stress preserves osmotic as well as ionic equilibria and maintains homeostasis (Dametto *et al*., 2015). Also, it is well known that low temperatures activate ion channels and soluble sugars transporters, especially at the chloroplast and tonoplast (Tarkowski & Van den Ende, 2015; Pommerrenig *et al*., 2018).

In contrast with studies performed in model species and plants from temperate environments, in our study, the majority of the DEG that possess GO annotation terms were associated with broad categories such as biological or metabolic processes, with just one exception associated with abiotic stress stimulus (GO:0009628). However, more than a half of the genes found in the freezing survival QTL interval were differentially expressed in the coldest treatment and all of them were associated with a known response to cold stress (Table 2.5). For instance, we discovered a *P. hallii* gene from the zinc finger family which is expressed during cold stress in tobacco and rice (Mukhopadhyay *et al*., 2004; Shakiba *et al*., 2017). Similarly, we identified a gene involved in energy metabolism, the Cytochrome C oxidase copper chaperone (COX17), that was also expressed in loquat (*Eriobotrya japonica*) under freezing stress (Xu *et al*., 2017). In addition, the DEAD-box family gene that was up-regulated in our experiment and has previously observed to be involved in cold responses in *A. thaliana* (Gong *et al*., 2004; Chinnusamy *et al*., 2007). Lastly, a sulfite exporter TauE/SafE family protein was down-regulated and has previously

been detected in response to freezing temperatures in wine grape (*Vitis vinifera*). These genes are promising candidates underlying freezing tolerance in *Panicum hallii*.

Conclusion

Here, we extend previous studies of cold-stress gene expression in plants to incorporate a more realistic stress treatment in C_4 perennial grass. In addition, we utilize QTL and global gene expression approaches using 3'-TaqSeq. Using these approaches, it was possible to describe the early chilling responses of *P. hallii* and identify11 candidate genes that may underlie freezing tolerance. Accordingly, the results here not only aid in the discovery of the genetic mechanism that underline local adaptation but also provide a foundation to improve switchgrass yield under cold condition. Future studies at the transcriptome level should include an extended period of time over freezing temperatures along with sampling across recovery phases. These data should help to elucidate early signaling components as well evaluate mechanisms of recovering in perennial plants. Finally, careful studies of the 11 DEG found in the survival QTL might provide new insights about how plants can deal with low temperatures.

TABLES

Table 2.1 Means, one standard error (SE) and broad-sense heritability (H^2) of freezing tolerance and regrowth traits in the *Panicum hallii* RIL population and its parents.

Phenotypic trait	190 RILs (4 replicates)	FIL2 (n=33)	HAL2 (n=33)	*P-value*[1]	H^2 ±SE
Survival after 1-month (Proportion)	0.28±0.45	0.45±0.08	0.21±0.08	0.0371	0.11±0.03
Tiller number 1-month (Count)	2.96±2.04	4.73±0.48	2.71±0.7	0.0290	
Survival after 3-months (Proportion)	0.26±0.44	0.45±0.07	0.18±0.18	0.0171	0.9±0.04
Tiller number after 3-months (Count)	2.54±1.47	4.13±0.47	2.5±0.74	0.0794	

[1] *P-values* correspond to F-test from ANOVA contrasting parents (FIL2 versus HAL2)

Table 2.2 QTL identified for freezing tolerance and regrowth in *Panicum hallii* RIL population. QTL confidence intervals were determined by a 1.5 LOD drop from the QTL LOD peak.

Phenotypic trait	Chromosome	Peak (cM)	LOD	Confidence interval (cM)	% variance
Survival 1-month	9	17.26	3.16	14.24-43.56	7.26
Tiller number 1-month	9	38.01	3.37	14.24-72.13	6.86
Tiller number 3-months	9	26.46	3.73	14.24-47.96	8.52

Table 2.3 Differentially expressed genes by temperature treatment at different time intervals across the day as identified using DEseq contrasts and an FDR threshold of 0.05.

Treatment	Total genes	Up regulated	Down regulated	Exclusive genes
26 versus 21ºC/ZT8	1,208	464 (38.4%)	744 (61.6%)	238 (19.7%)
26 versus 15ºC/ZT14	8,598	4,268 (49.6%)	4,330 (50.4%)	3,451 (40.13%)
26 versus 9ºC/ZT20	1,751	687 (39.2%)	1,064 (60.8%)	386 (22%)
26 versus 4ºC/ZT2	8,859	4,248 (48%)	4,611 (52%)	3,616 (40.8%)

Table 2.4 Differentially expressed genes in the QTL intervals under genotype-by-treatment model as detected by DEseq contrasts and an FDR threshold of 0.05.

Trait	Chromosome	Genes in the interval	Genotype-by-treatment interaction genes (temperature (ºC)/ ZT)			
			21/ZT8	15/ZT14	9/ZT20	4/ZT2
Survival 1-month	9	883	2	0	3	6
Tiller number 1-month	9	1,539	5	1	3	19
Tiller number 3-months	9	974	2	0	3	10

Table 2.5 List of genes that show significant genotype-by-treatment interactions present in the three overlapping QTL intervals across the four different temperature/ZT combinations and associated response to cold tolerance reported in the literature.

Gene ID[1]	QTL for traits			Treatment	Low temperature associated response in the literature
	Sur.	Til1	Til2		
Pahal.I01655	X	X	X	4ºC/ZT2	Carbohydrate-metabolism-related protein, down-regulated in cold acclimation in Arabidopsis (Oono *et al*., 2006).
Pahal.I02121	X	X	X	4ºC/ZT2	Involved in the process of spliceosome assembly of cold tolerance genes in hardy corylus species (*C. heterophylla* Fisch) floral buds (Chen *et al*., 2014).
Pahal.I02564		X	X	4ºC/ZT2	Arabidopsis CBF regulon genes and identification of putative tomato homologs and their expression in response to CBF overexpression (Carvallo *et al*., 2011).
Pahal.I02582		X	X	4ºC/ZT2	GsGASA1 mediated root growth inhibition in response to chronic cold stress is marked by the accumulation of DELLAs (Li *et al*., 2011).
Pahal.I02638		X	X	4ºC/ZT2	Transcriptome profiling of *Eucalyptus nitens* reveals deeper insight into the molecular mechanism of cold acclimation and deacclimation process (Gaete-Loyola *et al*., 2017).
Pahal.I02662		X	X	4ºC/ZT2	microRNAs and Their Targets in Cold-Stored Potato Tubers (Ou *et al*., 2015).
Pahal.I02758		X		4ºC/ZT2	Unannotated
Pahal.I02960		X		4ºC/ZT2	Early down-regulation of an alpha-galactosidase in barley response to cold (Janská *et al*., 2011).
Pahal.I03210		X		4ºC/ZT2	Genotype-dependent Burst of Transposable Element Expression in Crowns of Hexaploid Wheat (Triticum aestivum L) during Cold Acclimation (Kim *et al*., 2014).
Pahal.I03888		X		4ºC/ZT2	Differentially expressed genes identified by RNAseq in rice germinating seeds of cold-tolerant and cold-sensitive genotypes (Dametto *et al*., 2015).
Pahal.J01443		X		4ºC/ZT2	Unannotated
Pahal.J01447		X		4ºC/ZT2	Elevated annotated as a DUF3506 domain-containing protein, shares strong similarity with the EX1 and EX2 proteins from A. thaliana proteins that detoxify light-stress generated singlet oxygen (Burns *et al*., 2018).
Pahal.I02059	X	X	X	4ºC/ZT2	Photosynthesis: proteins involved in G signaling pathways, including two Ran BP2/NZF zinc finger-like superfamily proteins (Ning *et al*., 2016).
Pahal.I02457	X	X	X	4ºC/ZT2	The transcript abundance of cytochrome c oxidase COX17 was increased in freezing treatment in *Eriobotrya joponica* (Xu *et al*., 2017).

Table 2.5 (Continue)

Gene ID[1]	QTL for traits			Treatment	Low temperature associated response in the literature
Pahal.I01899	X	X	X	4ºC/ZT2	The transduction of environmental signals into the nucleus to alter transcription and the export of mRNAs and small RNAs to the cytoplasm through the nuclear pore complex (NPC) of the nuclear envelop are crucial to gene regulation in eukaryotes (Chinnusamy *et al*., 2007).
Pahal.I01918	X	X	X	4 ºC/ZT2	Associated with transcripts with significantly different expression in VvCBF4 overexpressor in grapes (Tillett *et al*., 2012).
Pahal.I02697		X		4 ºC/ZT2	Unannotated
Pahal.I03118		X		4 ºC/ZT2	In Saccharomyces cerevisiae, the osmoregulatory pathways begin with either an Src-homology 3 (SH3)-domain-containing membrane protein or a two-component histidine kinase, which activates a mitogen-activated protein kinase (MAPK) cascade and leads to increased osmolyte synthesis and accumulation (Zhu, 2001).
Pahal.J01433		X		4 ºC/ZT2	Expression is relative to yellow-leaf-specific protein8 (YLS8) (Purdy *et al*., 2011).
Pahal.I01784	X	X	X	9 ºC/ZT20	Unannotated
Pahal.I01805	X	X	X	9 ºC/ZT20	Unannotated
Pahal.I01927	X	X	X	9 ºC/ZT20	Unannotated
Pahal.I03019		X		15 ºC/ZT14	higher cell membrane stability during the cold stress in rice (Chinnusamy *et al*., 2007; Hu *et al*., 2008).
Pahal.I01704	X	X	X	21 ºC/ZT8	Vernalization gene and photoperiod, K-box and MADS-box transcription factor family protein (Suo *et al*., 2016; Amasino, 2018)
Pahal.I02795		X		21 ºC/ZT8	unannotated
Pahal.I01810	X	X	X	21 ºC/ZT8	The 50 most highly expressed genes in 1 h cold-treated leaves (Byun *et al*., 2009).
Pahal.I02946		X		21 ºC/ZT8	A. thaliana REIL paralogs required at two check-23 points of leaf development in the cold (Schmidt *et al*., 2013).
Pahal.I02998		X		21 ºC/ZT8	Downregulated protein degradation (Seki *et al*., 2002; Oono *et al*., 2006).

[1]The *Panicum hallii* variety *filipes* v. 2 gene annotation (Phytozome ID). Sur.: survival after a month. Til1: Tiller number after a month. Til2: Tiller number after three months.

FIGURES

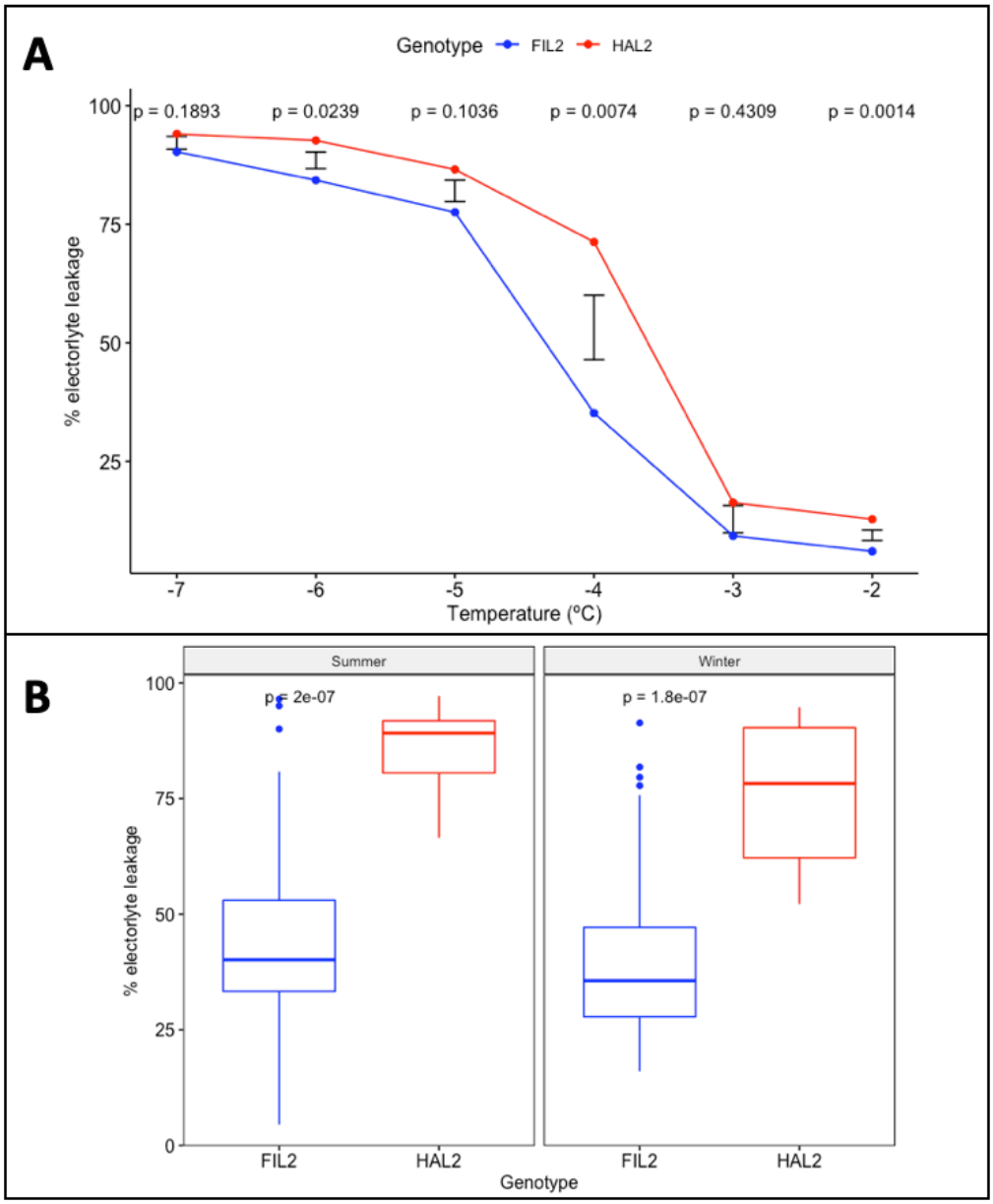

Figure 2.1. Percentage of relative conductivity as response to cold treatment in the parents of the RIL population. A. Relative conductivity response to temperatures between -7 to -2 ºC. The major significative difference between genotypes was scored at -4 ºC. B. Electrolyte leakage of plants of FIL2 and HAL2 growing under two different conditions simulating average temperatures in Austin TX during Summer and Winter.

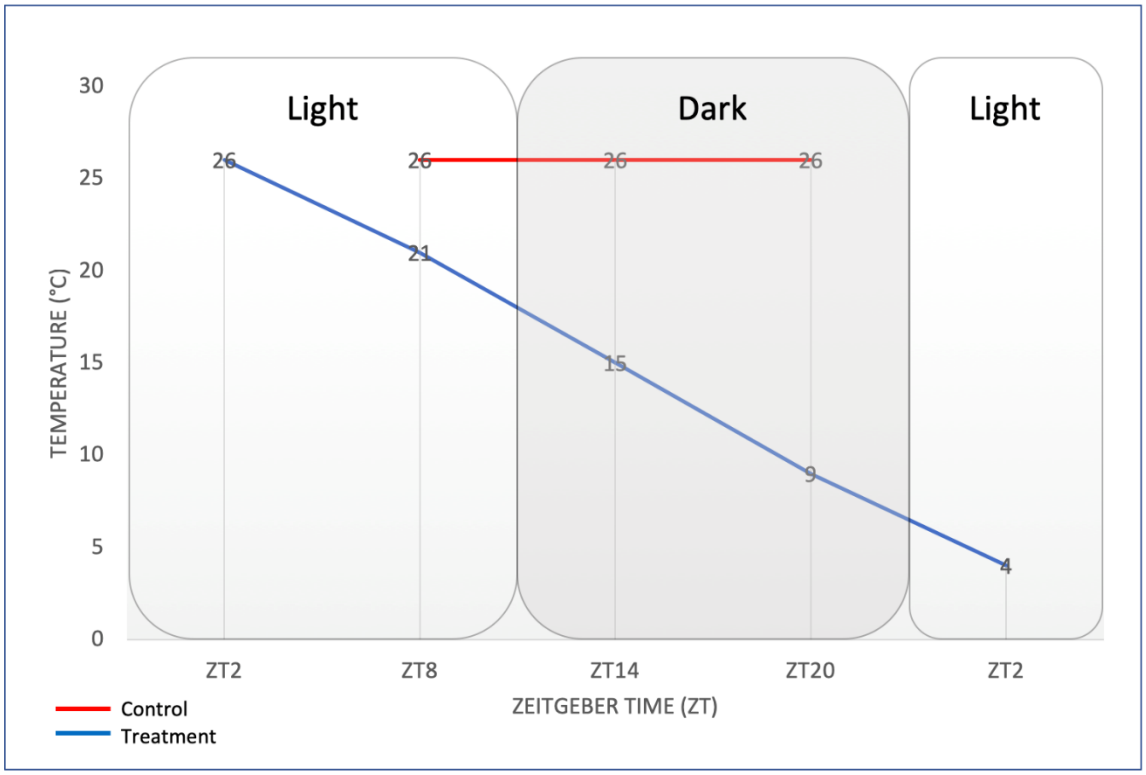

Figure 2.2. Graphical depiction of temperature treatments imposed across a temporal series that forms the basis of our studies of gene expression responses to chilling. The blue line represents treatment points while the red line represents the 26ºC control condition. The experimental period spanned both light and dark photoperiod conditions indicated by shading in the graphic and time showed as Zeitgeber time (ZT) along a circadian continuum. Sampling (n=3) were harvested from each genotype at time period/treatment combination except for ZT2 after 24 hrs where samples were only harvested at 4ºC.

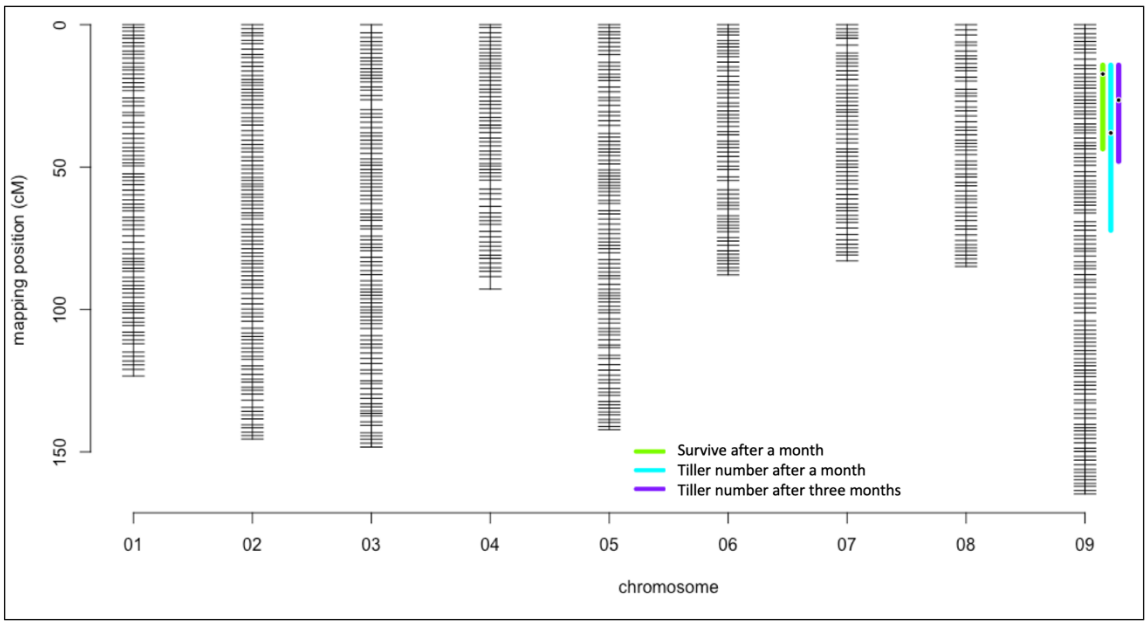

Figure 2.3. Genetic map of the location of the freezing response trait QTL in *Panicum hallii*. The length of colored bars indicates 1.5-LOD drop confidence intervals. The maximum LOD score for each QTL peak is presented by the dot within the bars.

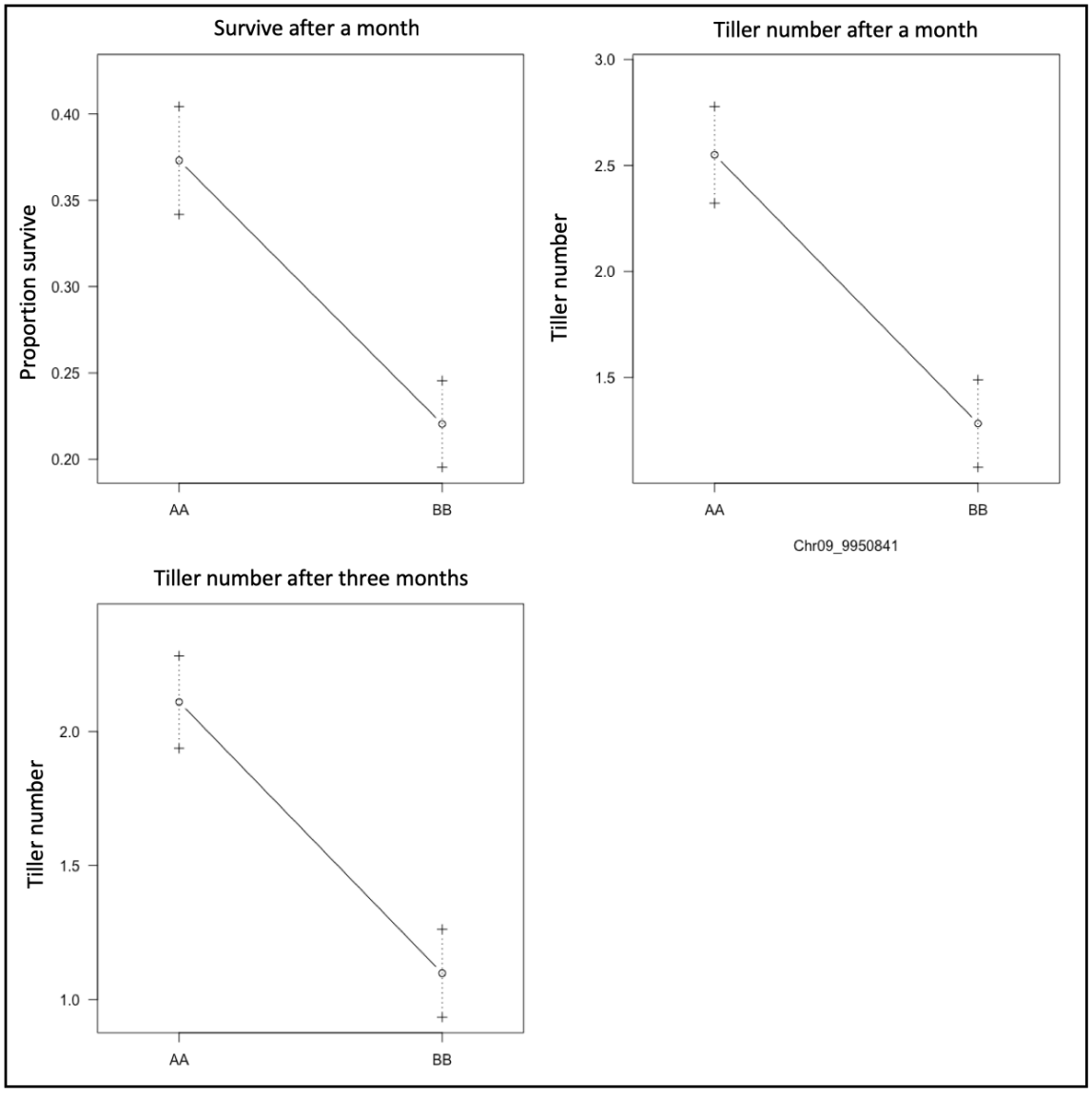

Figure 2.4. Additive effect plot for each QTL. AA genotype correspond to FIL2 parent and BB to HAL2 parent. It was calculated using the function *effecplot* in the R package rqtl2. Error bars show the standard error.

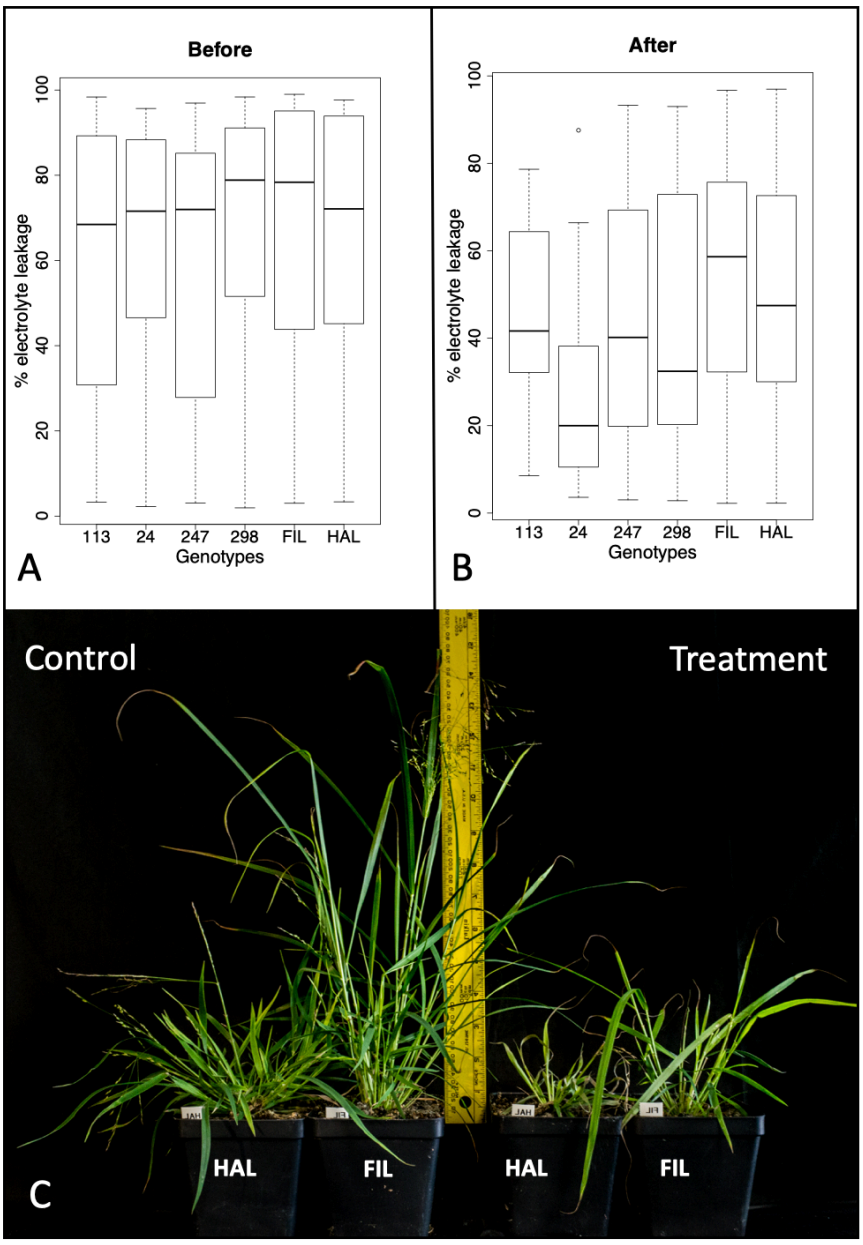

Figure 2.5. Phenotypic responses measure as electrolyte leakage. A. Boxplot of relative conductivity (RC) responses ranging from 0% (no damage; freezing tolerant) to 100% (complete damage; freezing sensitive) in all six genotypes included in the experiment in non-acclimatized plants (Mean = 62.46, SE±2.08) B. Acclimated plants after at 4ºC cold treatment (Mean = 44.72, SE±2.1). Treatment effects were significant (F= 37.89, P < .0001). C. Phenotypic appearance for HAL and FIL phenotypes between control and after treatment.

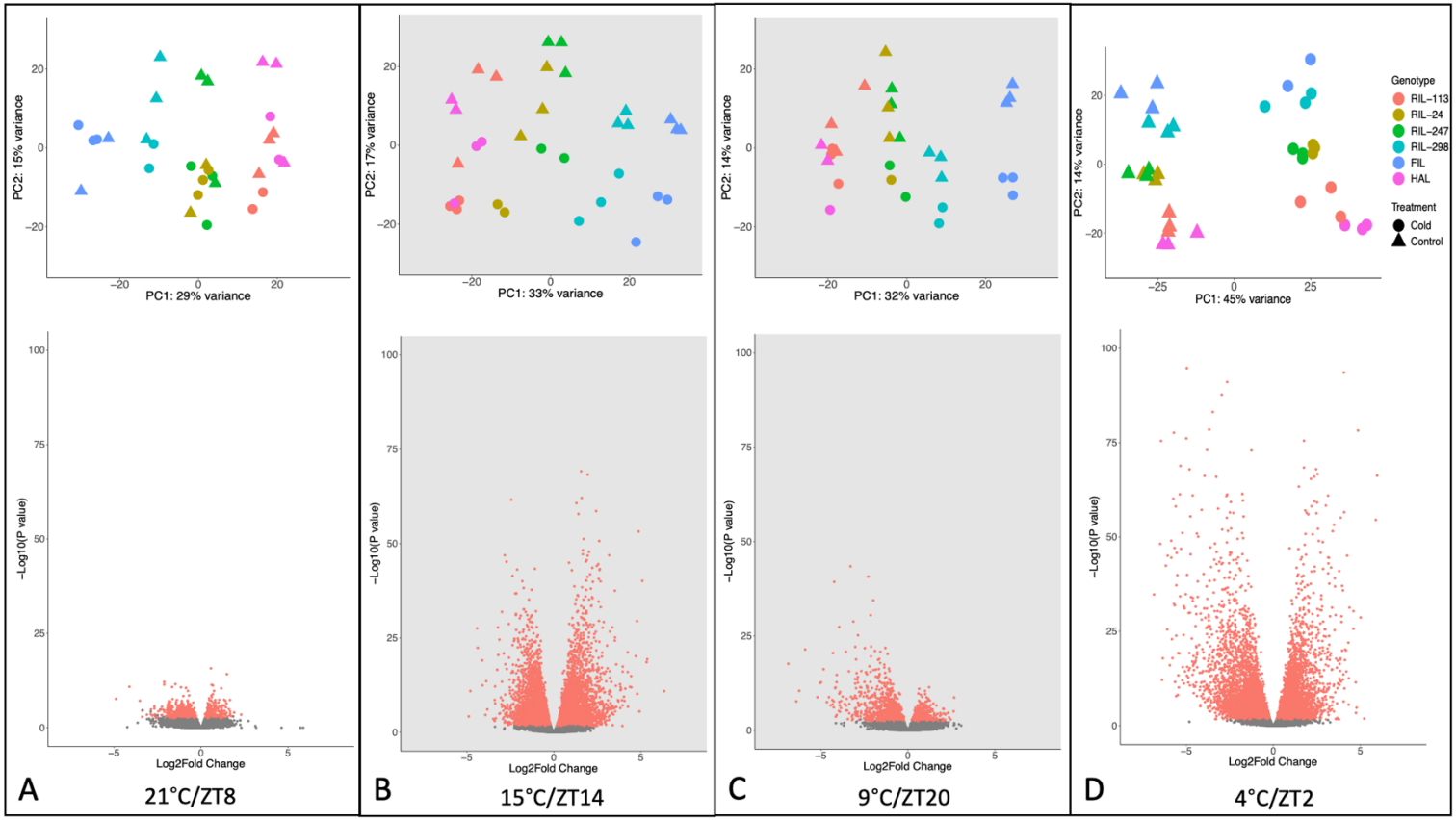

Figure 2.6. Differential gene expression responses to cold across four time points and temperature treatments. Principal component analysis of normalized data is plotted and grouped by treatments in the top. Axis legend include the percent variance explained by the first two PCA axes. Volcano plots at the bottom show the differential expression between treatments at FDR level of 0.05, where the log2 fold change of treatments compared to control is plotted on the horizontal axis and the *P*-value of the associated test is on the vertical. Shaded areas indicate dark photoperiod conditions.

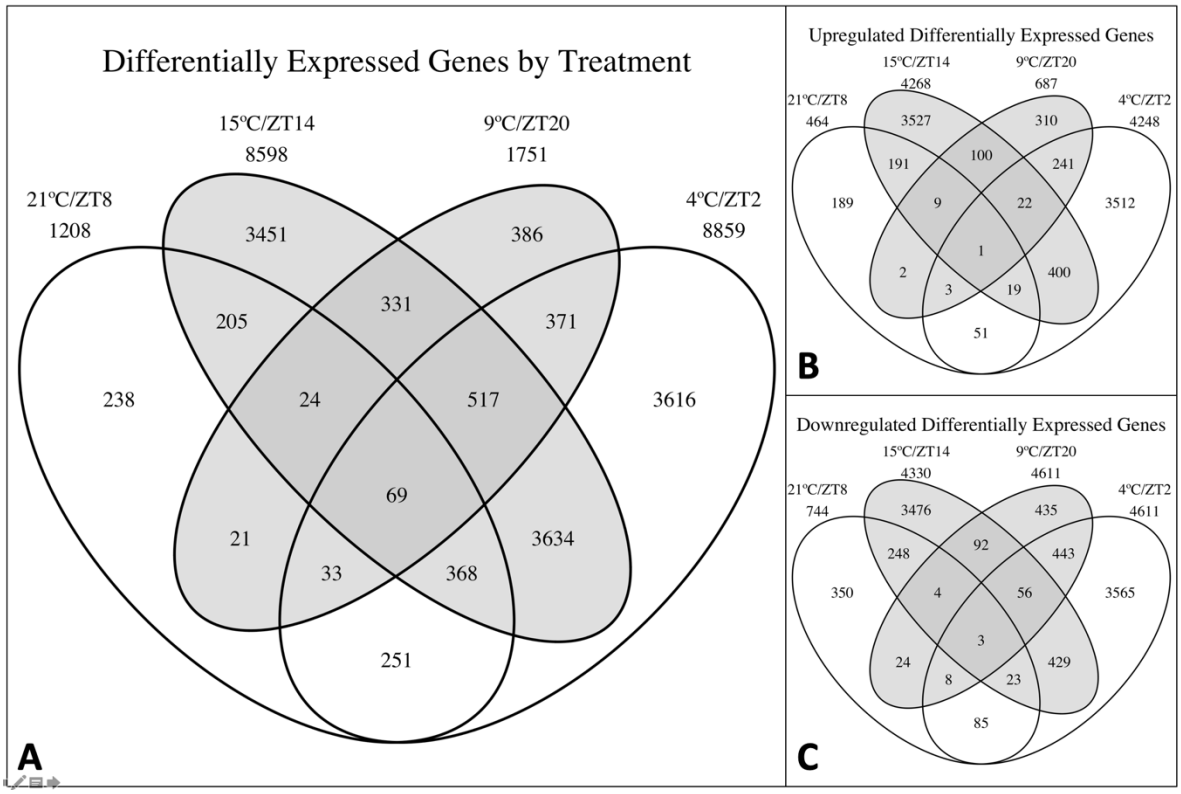

Figure 2.7. Venn diagrams with A. All differentially expressed genes in *Panicum hallii* discovered during a time series of decreasing temperature treatments, B. Upregulated differentially expressed genes, and, C. Downregulated differentially expressed genes. Treatment time and temperature, as well as the number of genes are indicated adjacent to the diagram. Shaded areas indicate dark photoperiod conditions.

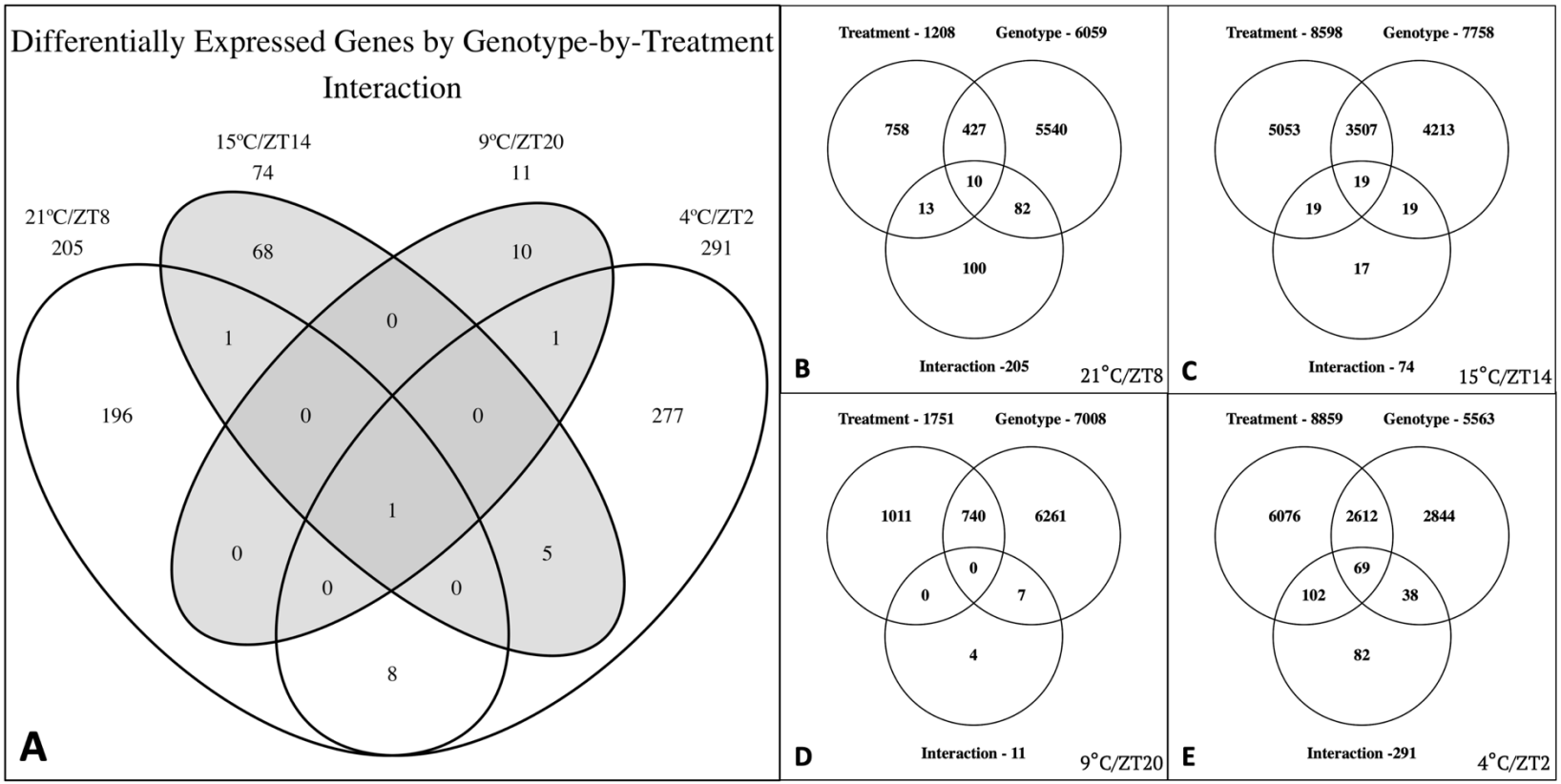

Figure 2.8. Venn diagram representing the overlap of gene list for (A) all the treatment by genotype interaction for each treatment and (B-E) experimental factors (Treatment, genotype, treatment by genotype) for each treatment.

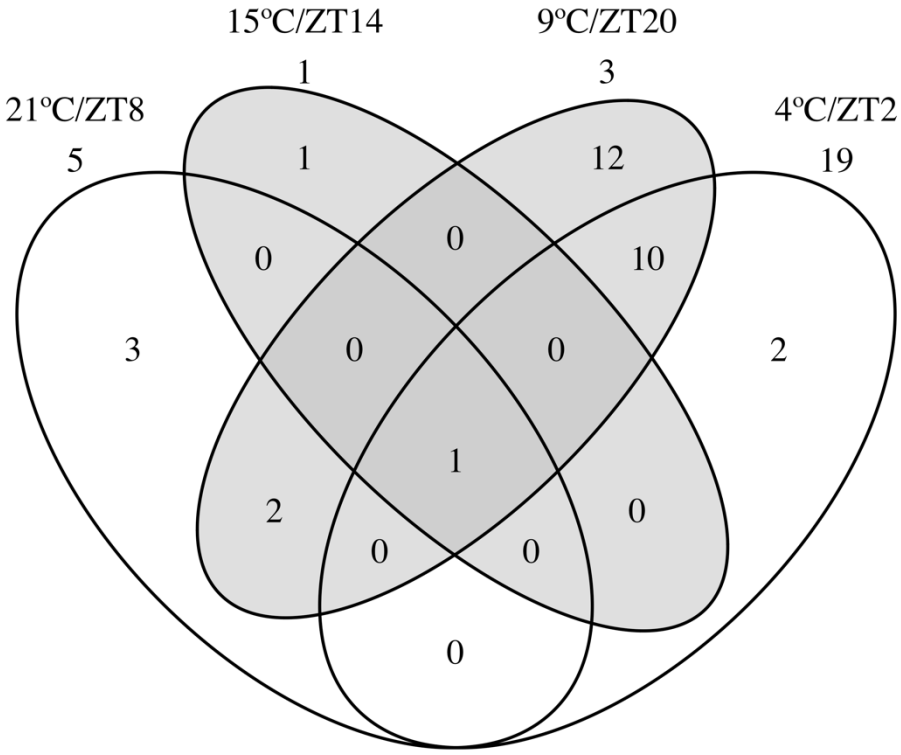

Figure 2.9. Venn diagram showing the genes discovered for each temperature/ZT treatment point present in the QTL intervals. Shaded areas indicate dark photoperiod conditions.

References

Alexa A, Ranhnenfuhrer J. 2016. topGo: Enrichment Analysis for Gene Ontology.

Allen GC, Flores-Vergara MA, Krasynanski S, Kumar S, Thompson WF. 2006. A modified protocol for rapid DNA isolation from plant tissues using cetyltrimethylammonium bromide. *Nature Protocols* **1**: 2320–2325.

Amasino R. 2018. A path to a biennial life history. *Nature Plants* **4**: 752–753.

Andrews S. 2018. FastQC: a quality control tool for high throughput sequence data.

Auton A, Abecasis GR, Altshuler DM, Durbin RM, Bentley DR, Chakravarti A, Clark AG, Donnelly P, Eichler EE, Flicek P, *et al.* 2015. A global reference for human genetic variation. *Nature* **526**: 68–74.

Baker HG. 1972. Seed Weight in Relation to Environmental Conditions in California. *Ecology* **53**: 997–1010.

Bryson RRW, Prendini L, Savary WEW, Pearman PB, Hewitt G, Avise J, Hewitt G, Webb T, Bartlein P, Hewitt G, *et al.* 2014. Caves as microrefugia: Pleistocene phylogeography of the troglophilic North American scorpion Pseudouroctonus reddelli. *BMC evolutionary biology* **14**: 9.

Burns EE, Keith BK, Refai MY, Bothner B, Dyer WE. 2018. Constitutive redox and phosphoproteome changes in multiple herbicide resistant Avena fatua L. are similar to those of systemic acquired resistance and systemic acquired acclimation. *Journal of Plant Physiology* **220**: 105–114.

Bushnell B. 2014. *BBMap: a fast, accurate, splice-aware aligner*. Ernest Orlando Lawrence Berkeley National Laboratory, Berkeley, CA (US).

Byun YJ, Kim HJ, Lee DH. 2009. LongSAGE analysis of the early response to cold stress in Arabidopsis leaf. *Planta* **229**: 1181–1200.

Campitelli BE, Gorton AJ, Ostevik KL, Stinchcombe JR. 2013. The effect of leaf shape on the thermoregulation and frost tolerance of an annual vine, Ipomoea hederacea (Convolvulaceae). *American Journal of Botany* **100**: 2175–2182.

Carvallo MA, Pino MT, Jeknić Z, Zou C, Doherty CJ, Shiu SH, Chen THH, Thomashow MF. 2011. A comparison of the low temperature transcriptomes and CBF regulons of three plant species that differ in freezing tolerance: Solanum commersonii, Solanum tuberosum, and Arabidopsis thaliana. *Journal of Experimental Botany* **62**: 3807–3819.

Casler MD, Tobias CM, Kaeppler SM, Buell CR, Wang Z-Y, Cao P, Schmutz J, Ronald P. 2011. The Switchgrass Genome: Tools and Strategies. *The Plant Genome*

Journal **4**: 273.

Casler M, Vogel K. **2004**. Latitudinal adaptation of switchgrass populations. *Crop Science*: 293–303.

Casler MD, Vogel KP. **2014**. Selection for biomass yield in upland, lowland, and hybrid switchgrass. *Crop Science* **54**: 626–636.

Casler MD, Vogel KP, Taliaferro CM, Ehlke NJ, Berdahl JD, Brummer EC, Kallenbach RL, West CP, Mitchell RB. **2007**. Latitudinal and longitudinal adaptation of switchgrass populations. *Crop Science* **47**: 2249–2260.

Catchen JM, Amores A, Hohenlohe P, Cresko W, Postlethwait JH. **2011**. Stacks: building and genotyping goci de novo from short-read sequences. *G3: Genes, Genomes, Genetics* **1**: 171–182.

Catchen J, Hohenlohe P a, Bassham S, Amores A, Cresko W a. **2013**. Stacks: an analysis tool set for population genomics. *Molecular ecology* **22**: 3124–40.

Chen H, Boutros PC. **2011**. VennDiagram: a package for the generation of highly-customizable Venn and Euler diagrams in R. *BMC bioinformatics* **12**: 35.

Chen X, Zhang J, Liu Q, Guo W, Zhao T, Ma Q, Wang G. **2014**. Transcriptome sequencing and identification of cold tolerance genes in hardy Corylus species (C. heterophylla fisch) floral buds. *PLoS ONE* **9**.

Chinnusamy V, Zhu J, Zhu JK. **2007**. Cold stress regulation of gene expression in plants. *Trends in Plant Science* **12**: 444–451.

Cramer GR, Urano K, Delrot S, Pezzotti M, Shinozaki K. **2011**. Effects of abiotic stress on plants: A systems biology perspective. *BMC Plant Biology* **11**: 163.

Dametto A, Sperotto RA, Adamski JM, Blasi ÉAR, Cargnelutti D, de Oliveira LFV, Ricachenevsky FK, Fregonezi JN, Mariath JEA, da Cruz RP, *et al.* 2015. Cold tolerance in rice germinating seeds revealed by deep RNAseq analysis of contrasting indica genotypes. *Plant Science* **238**: 1–12.

Eaton D a R. **2014**. PyRAD: Assembly of de novo RADseq loci for phylogenetic analyses. *Bioinformatics* **30**: 1844–1849.

Eaton DAR, Overcast I. **2016**. ipyrad: interactive assembly and analysis of RADseq data sets.

Elias SA. **1992**. Late Quaternary Zoogeography of the Chihuahuan Desert Insect Fauna, Based on Fossil Records from Packrat Middens. *Journal of Biogeography* **19**: 285.

Ellegren H, Galtier N. **2016**. Determinants of genetic diversity. *Nature Reviews Genetics* **17**: 422.

Estrada E, Villareal JÁ. **2010**. Flora del centro del estado de Chihuahua, México. *Acta Botanica Mexicana* **92**: 51–118.

da Fonseca RR, Albrechtsen A, Themudo GE, Ramos-Madrigal J, Sibbesen JA, Maretty L, Zepeda-Mendoza ML, Campos PF, Heller R, Pereira RJ. **2016**. Next-generation biology: Sequencing and data analysis approaches for non-model organisms. *Marine Genomics* **30**: 3–13.

Fowler S, Thomashow MF. **2002**. Arabidopsis Transcriptome Profiling Indicates That Multiple Regulatory Pathways Are Activated during Cold Acclimation in Addition to the CBF Cold Response Pathway. *The Plant Cell* **14**: 1675–1690.

Gaete-Loyola J, Lagos C, Beltrán MF, Valenzuela S, Emhart V, Fernández M. **2017**. Transcriptome profiling of Eucalyptus nitens reveals deeper insight into the molecular mechanism of cold acclimation and deacclimation process. *Tree Genetics and Genomes* **13**.

Gilmour SJ, Hajela RK, Thomashow MF. **1988**. Cold Acclimation in Arabidopsis thaliana. *Plant Physiology* **87**: 745–750.

Gong Z, Dong C, Lee H, Zhu J, Xiong L, Gong D. **2004**. A DEAD Box RNA Helicase Is Essential for mRNA Export and Important for Development and Stress Responses in Arabidopsis. *Plant Cell* **17**: 1–12.

Gould FW. **1975**. *The grasses of Texas*. Texas A&M University, Texas Agricultural Experiment Station.

Gould BA, Palacio-Mejia JD, Jenkins J, Mamidi S, Barry K, Schmutz J, Juenger TE, Lowry DB. **2018**. Population genomics and climate adaptation of a C4 perennial grass, Panicum hallii (Poaceae). *BMC Genomics* **19**: 792.

Haley CS, Knott SA. **1992**. A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. *Heredity* **69**: 315.

Hamrick JL, Godt MJW. **1996**. Effects of life history traits on genetic diversity in plant species. *Phil. Trans. R. Soc. Lond. B* **351**: 1291–1298.

He X, Sambe MAN, Zhuo C, Tu Q, Guo Z. **2015**. A temperature induced lipocalin gene from Medicago falcata (MfTIL1) confers tolerance to cold and oxidative stress. *Plant molecular biology* **87**: 645–54.

Hereford J. **2009**. A Quantitative Survey of Local Adaptation and Fitness Trade-Offs. *The American Naturalist* **173**: 579–588.

Herrera-Arrieta Y, Cortés-Ortiz A. **2010**. Listado florístico y aspectos ecológicos de la familia Poaceae para Chihuahua, Durango y Zacatecas, México. *Journal of the botanical research institute of Texas* **4**: 711–738.

Herrera AY, Pámanes DS. **2006**. Guía de pastos para el ganadero del Estado de Durango. *IPN-CONACYT-Fundación PRODUCE-Filo de Agua, Durango, México*.

Hewitt G. **2000**. The genetic legacy of the Quaternary ice ages. *Nature* **405**: 907–913.

Hewitt GM. **2001**. Speciation, hybrid zones and phylogeography--or seeing genes in space and time. *Molecular Ecology* **10**: 537–549.

Hewitt G. **2004**. Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society B: Biological Sciences* **359**: 183–195.

Holmgren CA, Norris J, Betancourt JL. **2007**. Inferences about winter temperatures and summer rains from the late Quaternary record of C4 perennial grasses and C3 desert shrubs in the northern Chihuahuan Desert. *Journal of Quaternary Science* **22**: 141–161.

Hu H, You J, Fang Y, Zhu X, Qi Z, Xiong L. **2008**. Characterization of transcription factor gene SNAC2 conferring cold and salt tolerance in rice. *Plant Molecular Biology* **67**: 169–181.

Humphreys M, Gasior D. **2013**. Cold tolerance. In: Kole C, ed. Genomics and breeding for climate-resilient crops. Berlin, Heidelberg: Springer Berlin Heidelberg, 133–165.

Janská A, Aprile A, Zámečník J, Cattivelli L, Ovesná J. **2011**. Transcriptional responses of winter barley to cold indicate nucleosome remodelling as a specific feature of crown tissues. *Functional & integrative genomics* **11**: 307–325.

Jombart T. **2008**. Adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics* **24**: 1403–1405.

Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, *et al.* 2012. Geneious. *Bioinformatics* **28**: 1647–9.

Khasanova A, Lovell JT, Bonnette JE, Jenkins J, Yoshinaga Y, Schmutz J, Juenger TE. **2018**. The genetic architecture of shoot and root trait divergence between upland and lowland ecotypes of a perennial grass. *bioRxiv*: 301531.

Kim D-H, Doyle MR, Sung S, Amasino RM. **2009**. Vernalization: Winter and the Timing of Flowering in Plants. *Annual Review of Cell and Developmental Biology* **25**: 277–299.

Kim SM, Suh JP, Lee CK, Lee JH, Kim YG, Jena KK. **2014**. QTL mapping and development of candidate gene-derived DNA markers associated with seedling cold tolerance in rice (Oryza sativa L.). *Molecular Genetics and Genomics* **289**: 333–343.

Knaus BJ, Grünwald NJ. **2017**. vcfr: A package to manipulate and visualize variant call format data in r. *Molecular Ecology Resources* **17**: 44–53.

Lee DK, Aberle E, Anderson EK, Anderson W, Baldwin BS, Baltensperger D,

Barrett M, Blumenthal J, Bonos S, Bouton J, *et al.* 2018. Biomass production of herbaceous energy crops in the United States: field trial results and yield potential maps from the multiyear regional feedstock partnership. *GCB Bioenergy* **10**: 698–716.

Lepais O, Weir JT. **2014**. SimRAD: a R package for simulation-based prediction of the number of loci expected in RADseq and similar genotyping by sequencing approaches. *Molecular Ecology Resources* **33**: 1314–1321.

Li KL, Bai X, Li Y, Cai H, Ji W, Tang LL, Wen YD, Zhu YM. **2011**. GsGASA1 mediated root growth inhibition in response to chronic cold stress is marked by the accumulation of DELLAs. *Journal of Plant Physiology* **168**: 2153–2160.

Li H, Durbin R. **2010**. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* **26**: 589–595.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. **2009**. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**: 2078–2079.

Li Z, Jiang D, He Y. **2018**. FRIGIDA establishes a local chromosomal environment for FLOWERING LOCUS C mRNA production. *Nature Plants* **4**: 836–846.

Li L, Mao D. **2018**. Deployment of cold tolerance loci from *Oryza sativa* ssp. *Japonica* cv. 'Nipponbare' in a high-yielding *Indica* rice cultivar '93-11'. *Plant Breeding* **137**: 553–560.

Li C, Rudi H, Stockinger EJ, Cheng H, Cao M, Fox SE, Mockler TC, Westereng B, Fjellheim S, Rognli OA, *et al.* 2012. Comparative analyses reveal potential uses of Brachypodium distachyon as a model for cold stress responses in temperate grasses. *BMC Plant Biology* **12**: 65.

Liao Y, Smyth GK, Shi W. **2013**. The Subread aligner: Fast, accurate and scalable read mapping by seed-and-vote. *Nucleic Acids Research* **41**.

Liao Y, Smyth GK, Shi W. **2014**. FeatureCounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**: 923–930.

Lohman BK, Weber JN, Bolnick DI. **2016**. Evaluation of TagSeq, a reliable low-cost alternative for RNAseq. *Molecular Ecology Resources* **16**: 1315–1321.

Love MI, Huber W, Anders S. **2014**. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* **15**: 550.

Lovell JT, Jenkins J, Lowry DB, Mamidi S, Sreedasyam A, Weng X, Barry K, Bonnette J, Campitelli B, Daum C, *et al.* 2018. The genomic landscape of molecular responses to natural drought stress in Panicum hallii. *Nature Communications* **9**: 5213.

Lovell JT, Shakirov E V, Schwartz S, Lowry DB, Aspinwall MJ, Taylor SH, Bonnette J, Palacio-Mejia JD, Hawkes C V., Fay PA, *et al.* 2016. Promises and challenges of eco-physiological genomics in the field: tests of drought responses in switchgrass. *Plant Physiology* **172**: 734–748.

Lowry DB, Hernandez K, Taylor SH, Meyer E, Logan TL, Barry KW, Chapman JA, Rokhsar DS, Schmutz J, Juenger TE. 2015. The genetics of divergence and reproductive isolation between ecotypes of Panicum hallii. *New Phytologist* **205**: 402–414.

Lowry DB, Modliszewski JL, Wright KM, Wu CA, Willis JH. 2008. The strength and genetic basis of reproductive isolating barriers in flowering plants. *Philosophical Transactions of the Royal Society B: Biological Sciences* **363**: 3009 LP-3021.

Lowry DB, Purmal CT, Juenger TE. 2013. A population genetic transect of Panicum hallii (Poaceae). *American Journal of Botany* **100**: 592–601.

Lowry DB, Purmal CT, Meyer E, Juenger TE. 2012. Microsatellite markers for the native Texas perennial grass, Panicum hallii (Poaceae). *American Journal of Botany* **99**: 114–116.

MacRoberts MH, Sorrie BA, MacRoberts BR, Evans RE. 2002. Endemism in the West Gulf Coastal Plain: importance of xeric habitats. *SIDA, Contributions to Botany* **20**: 767–780.

Majure LC, Judd WS, Soltis PS, Soltis DE. 2012. Cytogeography of the Humifusa clade of Opuntia s.s. Mill. 1754 (Cactaceae, Opuntioideae, Opuntieae): Correlations with pleistocene refugia and morphological traits in a polyploid complex. *Comparative Cytogenetics* **6**: 53–77.

Maleki M, Ghorbanpour M. 2018. *Cold Tolerance in Plants: Molecular Machinery Deciphered*. Elsevier Inc.

Des Marais DL, McKay JK, Richards JH, Sen S, Wayne T, Juenger TE. 2012. Physiological Genomics of Response to Soil Drying in Diverse Arabidopsis Accessions. *The Plant Cell* **24**: 893–914.

Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17**: 10.

Mastretta-Yanes A, Arrigo N, Alvarez N, Jorgensen TH, Piñero D, Emerson BC. 2015. Restriction site-associated DNA sequencing, genotyping error estimation and de novo assembly optimization for population genetic inference. *Molecular Ecology Resources* **15**: 28–41.

McClung CR. 2006. Plant Circadian Rhythms. *the Plant Cell Online* **18**: 792–803.

McClung CR. **2015**. Circadian clocks: Who knows where the time goes. *Nature Plants* **1**: 15172.

McKeown M, Schubert M, Marcussen T, Fjellheim S, Preston JC. **2016**. Evidence for an early origin of vernalization responsiveness in temperate Pooideae grasses. *Plant Physiology* **172**: 416–426.

McKinney GJ, Waples RK, Seeb LW, Seeb JE. **2017**. Paralogs are revealed by proportion of heterozygotes and deviations in read ratios in genotyping-by-sequencing data from natural populations. *Molecular Ecology Resources* **17**: 656–669.

Meyer E, Aglyamova G V., Matz M V. 2011. Profiling gene expression responses of coral larvae (Acropora millepora) to elevated temperature and settlement inducers using a novel RNA-Seq procedure. *Molecular Ecology* **20**: 3599–3616.

Meyer E, Aspinwall MJ, Lowry DB, Palacio-Mejía JD, Logan TL, Fay PA, Juenger TE. **2014**. Integrating transcriptional, metabolomic, and physiological responses to drought stress and recovery in switchgrass (Panicum virgatum L.). *BMC Genomics* **15**: 527.

Meyer E, Logan TL, Juenger TE. **2012**. Transcriptome analysis and gene expression atlas for Panicum hallii var. filipes, a diploid model for biofuel research. *The Plant journal : for cell and molecular biology* **70**: 879–90.

Milano ER, Lowry DB, Juenger TE. **2016**. The Genetic Basis of Upland/Lowland Ecotype Divergence in Switchgrass (Panicum virgatum). *G3: Genes|Genomes|Genetics* **6**: 3561 LP-3570.

Mora-Márquez F, García-Olivares V, Emerson BC, López de Heredia U. **2017**. ddradseqtools : a software package for in silico simulation and testing of double-digest RADseq experiments. *Molecular Ecology Resources* **17**: 230–246.

Mukhopadhyay A, Vij S, Tyagi AK. **2004**. Overexpression of a zinc-finger protein gene from rice confers tolerance to cold, dehydration, and salt stress in transgenic tobacco. *Proceedings of the National Academy of Sciences* **101**: 6309–6314.

Nguyen LT, Schmidt HA, Von Haeseler A, Minh BQ. **2015**. IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution* **32**: 268–274.

Ning DL, Liu KH, Liu CC, Liu JW, Qian CR, Yu Y, Wang YF, Wang YC, Wang BC. **2016**. Large-scale comparative phosphoprotein analysis of maize seedling leaves during greening. *Planta* **243**: 501–517.

Noss RF. **2014**. *Forgotten grasslands of the south: Natural history and conservation*. Island Press.

Noss RF, Platt WJ, Sorrie BA, Weakley AS, Means DB, Costanza J, Peet RK. **2015**. How global biodiversity hotspots may go unrecognized: Lessons from the North American Coastal Plain. *Diversity and Distributions* **21**: 236–244.

Olson DM, Dinerstein E, Wikramanayake ED, Burgess ND, Powell GVN, Underwood EC, D'amico JA, Itoua I, Strand HE, Morrison JC, *et al.* 2001. Terrestrial Ecoregions of the World: A New Map of Life on Earth. *BioScience* **51**: 933.

Oono Y, Seki M, Satou M, Iida K, Akiyama K, Sakurai T, Fujita M, Yamaguchi-Shinozaki K, Shinozaki K. **2006**. Monitoring expression profiles of Arabidopsis genes during cold acclimation and deacclimation using DNA microarrays. *Functional & integrative genomics* **6**: 212–234.

Ortiz EM. **2018**. paralog-finder: a script to detect and blacklist paralog RAD loci.

Ou Y, Liu X, Xie C, Zhang H, Lin Y, Li M, Song B, Liu J. **2015**. Genome-wide identification of microRNAs and their targets in cold-stored potato tubers by deep sequencing and Degradome analysis. *Plant Molecular Biology Reporter* **33**: 584–597.

Ozsolak F, Milos PM. **2011**. RNA sequencing: advances, challenges and opportunities. *Nature reviews. Genetics* **12**: 87–98.

Pandit A, Rai V, Bal S, Sinha S, Kumar V, Chauhan M, Gautam RK, Singh R, Sharma PC, Singh AK, *et al.* 2010. Combining QTL mapping and transcriptome profiling of bulked RILs for identification of functional polymorphism for salt tolerance genes in rice (Oryza sativa L.). *Molecular Genetics and Genomics* **284**: 121–136.

Peixoto M de M, Sage RF. **2017**. Comparative photosynthetic responses in upland and lowland sugarcane cultivars grown in cool and warm conditions. *Brazilian Journal of Botany* **40**: 829–839.

Pérez-Harguindeguy N, Díaz S, Garnier É, Lavorel S, Poorter H, Jaureguiberry P, Bret-Harte MS, Cornwell WK, Craine JM, Gurvich DE, *et al.* 2013. New handbook for standardised measurement of plant functional traits worldwide. *Australian Journal of Botany* **61**: 167–234.

Peterson BK, Weber JN, Kay EH, Fisher HS, Hoekstra HE. **2012**. Double Digest RADseq: An Inexpensive Method for De Novo SNP Discovery and Genotyping in Model and Non-Model Species (L Orlando, Ed.). *PLoS ONE* **7**: e37135.

Phillips PC, Arnold SJ. **1999**. Hierarchical Comparison of Genetic Variance-Covariance Matrices. I. Using the Flury Hierarchy. *Evolution* **53**: 1506–1515.

Pommerrenig B, Ludewig F, Cvetkovic J, Trentmann O, Klemens PAW, Neuhaus HE. **2018**. In concert: Orchestrated changes in carbohydrate homeostasis are critical for plant abiotic stress tolerance. *Plant and Cell Physiology* **59**: 1290–1299.

Porras-Hurtado L, Ruiz Y, Santos C, Phillips C, Carracedo Á, Lareu M V. 2013. An overview of STRUCTURE: Applications, parameter settings, and supporting software. *Frontiers in Genetics* **4**: 98.

Preston JC, Sandve SR. 2013. Adaptation to seasonality and the winter freeze. *Frontiers in plant science* **4**: 1–18.

Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus genotype data. *Genetics* **155**: 945–59.

Purdy SJ, Bussell JD, Nelson DC, Villadsen D, Smith SM. 2011. A nuclear-localized protein, KOLD SENSITIV-1, affects the expression of cold-responsive genes during prolonged chilling in Arabidopsis. *Journal of Plant Physiology* **168**: 263–269.

R Core Team. 2018. R: A language and environment for statistical computing. R Foundation for Statistical Computing.

Razgour O, Juste J, Ibáñez C, Kiefer A, Rebelo H, Puechmaille SJ, Arlettaz R, Burke T, Dawson DA, Beaumont M, *et al.* 2013. The shaping of genetic variation in edge-of-range populations under past and future climate change. *Ecology Letters* **16**: 1258–1266.

Rebernig CA, Schneeweiss GM, Bardy KE, Schönswetter P, Villaseñor JL, Obermayer R, Stuessy TF, Weiss-Schneeweiss H. 2010. Multiple pleistocene refugia and holocene range expansion of an abundant southwestern American desert plant species (Melampodium leucanthum, Asteraceae). *Molecular Ecology* **19**: 3421–3443.

Reid CS, Urbatsch L. 2012. Noteworthy plant records from Louisiana. *Journal of the Botanical Research Institute of Texas* **6**: 273–278.

Remington CL. 1968. Suture-zones of hybrid interaction between recently joined biotas. In: Evolutionary biology. Springer, 321–428.

Robertson GP, Hamilton SK, Barham BL, Dale BE, Izaurralde RC, Jackson RD, Landis DA, Swinton SM, Thelen KD, Tiedje JM. 2017. Cellulosic biofuel contributions to a sustainable energy future: Choices and outcomes. *Science* **356**: 1349.

Rochette NC, Catchen JM. 2017. Deriving genotypes from RAD-seq short-read data using Stacks. *Nature Protocols* **12**: 2640–2659.

Rodríguez VM, Butrón A, Rady MOA, Soengas P, Revilla P. 2014. Identification of quantitative trait loci involved in the response to cold stress in maize (Zea mays L.). *Molecular Breeding* **33**: 363–371.

Sánchez-Ken JG. 2011. *Flora del valle de Tehuacán-Cuicatlán223* (R Medina Lermos, Ed.). México, D.F.: UNAM.

Sandve SR, Kosmala A, Rudi H, Fjellheim S, Rapacz M, Yamada T, Rognli OA. **2011**. Molecular mechanisms underlying frost tolerance in perennial grasses adapted to cold climates. *Plant Science* **180**: 69–77.

Schmidt S, Dethloff F, Beine-Golovchuk O, Kopka J. **2013**. The REIL1 and REIL2 Proteins of Arabidopsis thaliana Are Required for Leaf Growth in the Cold. *Plant Physiology* **163**: 1623–1639.

Schubert M, Groenvold L, Sandve SR, Hvidsten TR, Fjellheim S. **2017**. Evolution of cold acclimation in temperate grasses (Pooideae). *bioRxiv*.

Seal JN, Brown L, Ontiveros C, Thiebaud J, Mueller UG. **2015**. Gone to Texas: comparative phylogeography of two Trachymyrmex species along the southeastern coastal plain of North America. *Biological Journal of the Linnean Society* **114**: 689–698.

Seki M, Narusaka M, Ishida J, Nanjo T, Fujita M, Oono Y, Kamiya A, Nakajima M, Enju A, Sakurai T, *et al.* 2002. Monitoring the expression profiles of 7000 Arabidopsis genes under drought, cold and high-salinity stresses using a full-length cDNA microarray. *Plant Journal* **31**: 279–292.

Shafer ABA, Peart CR, Tusso S, Maayan I, Brelsford A, Wheat CW, Wolf JBW. **2017**. Bioinformatic processing of RAD-seq data dramatically impacts downstream population genetic inference (M Gilbert, Ed.). *Methods in Ecology and Evolution* **8**: 907–917.

Shakiba E, Edwards JD, Jodari F, Duke SE, Baldo AM, Korniliev P, McCouch SR, Eizenga GC. **2017**. Genetic architecture of cold tolerance in rice (Oryza sativa) determined through high resolution genome-wide analysis (SK Parida, Ed.). *PLOS ONE* **12**: e0172133.

Shi Y, Ding Y, Yang S. **2015**. Cold signal transduction and its interplay with phytohormones during cold acclimation. *Plant and Cell Physiology* **56**: 7–15.

Shi Y, Ding Y, Yang S. **2018**. Molecular Regulation of CBF Signaling in Cold Acclimation. *Trends in Plant Science* **23**: 623–637.

Shorthouse D. **2018**. SimpleMappr, an online tool to produce publication-quality point maps.

Sinha S, Kukreja B, Arora P, Sharma M, Pandey GK, Agarwal M, Chinnuasamy V. **2015**. The omics of cold stress responses in plants. In: Elucidation of Abiotic Stress Signaling in Plants. 143–194.

Sobel JM, Chen GF. **2014**. Unification of methods for estimating the strength of reproductive isolation. *Evolution* **68**: 1511–1522.

Sorrie BA, Weakley AS. **2001**. Coastal plain vascular plant endemics: phytogeographic

patterns. *Castanea*: 50–82.

Sperotto RA, de Araújo Junior AT, Adamski JM, Cargnelutti D, Ricachenevsky FK, de Oliveira BHN, da Cruz RP, dos Santos RP, da Silva LP, Fett JP. **2018**. Deep RNAseq indicates protective mechanisms of cold-tolerant indica rice plants during early vegetative stage. *Plant Cell Reports* **37**: 347–375.

Steponkus PL. **1984**. Injury and Cold Acclimation. *Annual Review of Plant Physiology* **35**: 543–584.

Stuessy TF, Weiss-Schneeweiss H, Keil DJ. **2004**. Diploid and polyploid cytotype distribution in Melampodium cinereum and M. leucanthum (Asteraceae, Heliantheae). *American Journal of Botany* **91**: 889–898.

Suo H, Lü J, Ma Q, Yang CY, Zhang XX, Meng X, Huang S, Nian H. **2016**. The AtDREB1A transcription factor up-regulates expression of a vernalization pathway gene, GmVRN1-like, delaying flowering in soybean. *Acta Physiologiae Plantarum* **38**: 1–10.

Swenson NG, Howard DJ. **2004**. Do suture zone exist? *Evolution* **58**: 2391–2397.

Swenson, Howard. **2005**. Clustering of Contact Zones, Hybrid Zones, and Phylogeographic Breaks in North America. *The American Naturalist* **166**: 581.

Tarkowski ŁP, Van den Ende W. **2015**. Cold tolerance triggered by soluble sugars: a multifaceted countermeasure. *Frontiers in Plant Science* **6**: 1–7.

Thomashow MF. **1999**. PLANT COLD ACCLIMATION: Freezing Tolerance Genes and Regulatory Mechanisms. *Annual review of plant physiology and plant molecular biology* **50**: 571–599.

Tillett RL, Wheatley MD, Tattersall EAR, Schlauch KA, Cramer GR, Cushman JC. **2012**. The Vitis vinifera C-repeat binding protein 4 (VvCBF4) transcriptional factor enhances freezing tolerance in wine grape. *Plant Biotechnology Journal* **10**: 105–124.

Tyler L, Lee SJ, Young ND, DeIulio GA, Benavente E, Reagon M, Sysopha J, Baldini RM, Troìa A, Hazen SP, *et al.* 2016. Population Structure in the Model Grass Is Highly Correlated with Flowering Differences across Broad Geographic Areas. *The Plant Genome* **9**: 1–20.

Valdés-Reyna J, Villaseñor JL, Encina-Domínguez JA, Ortiz E. **2015**. The grass family (poaceae) in Coahuila, Mexico: Diversity and distribution. *Botanical Sciences* **93**: 119–129.

Vellend M, Geber MA. **2005**. Connections between species diversity and genetic diversity. *Ecology Letters* **8**: 767–781.

Waller FR. **1974**. A New Combination in Panicum (Gramineae). *The Southwestern*

Naturalist **19**: 105.

Waller FR. **1976**. A biosystematic study of Panicum section Diffusa (Poaceae) in North America. **37**.

Woods DP, Amasino RM. **2015**. Dissecting the control of flowering time in grasses using Brachypodium distachyon. *Plant Genetics and Genomics: Crops and Models*: 1–26.

Woods DP, McKeown MA, Dong Y, Preston JC, Amasino RM. **2016**. Evolution of *VRN2*/*Ghd7*- Like Genes in Vernalization-Mediated Repression of Grass Flowering. *Plant Physiology* **170**: 2124–2135.

Xu S, Chong K. **2018**. Remembering winter through vernalisation. *Nature Plants* **4**: 997–1009.

Xu H xia, Li X ying, Chen J wei. **2017**. Comparative transcriptome profiling of freezing stress responses in loquat (Eriobotrya japonica) fruitlets. *Journal of Plant Research* **130**: 893–907.

Yang QS, Gao J, He W Di, Dou TX, Ding LJ, Wu JH, Li CY, Peng XX, Zhang S, Yi GJ. **2015**. Comparative transcriptomics analysis reveals difference of key gene expression between banana and plantain in response to cold stress. *BMC Genomics* **16**: 1–18.

Zhang Y, Ngu DW, Carvalho D, Liang Z, Qiu Y, Roston RL, Schnable JC. **2017**. Differentially Regulated Orthologs in Sorghum and the Subgenomes of Maize. *The Plant Cell* **29**: 1938–1951.

Zhang Y, Zalapa JE, Jakubowski AR, Price DL, Acharya A, Wei Y, Brummer EC, Kaeppler SM, Casler MD. **2011**. Post-glacial evolution of Panicum virgatum: centers of diversity and gene pools revealed by SSR markers and cpDNA sequences. *Genetica* **139**: 933–48.

Zhao C, Lang Z, Zhu JK. **2015**. Cold responsive gene transcription becomes more complex. *Trends in Plant Science* **20**: 466–468.

Zhong J, Robbett M, Poire A, Preston JC. **2018**. Successive evolutionary steps drove Pooideae grasses from tropical to temperate regions. *New Phytologist* **217**: 925–938.

Zhu JK. **2001**. Cell signaling under salt, water and cold stresses. *Current Opinion in Plant Biology* **4**: 401–406.

Vita

Juan Diego Palacio Mejía born in Medellín, Colombia. Interested in the nature and countryside life style, he went to the University to pursuit his undergraduate studies in Agronomy at the Universidad Nacional de Colombia, located in Medellín. During this time, he developed special interest in "orphan" crops such as some tropical fruits, vegetables and tubers and in biotechnology tools. As a result, he did his thesis in cryoconservation of cassava in the International Center of Tropical Agriculture. After he graduated, he got a job in the Humboldt Institute, the Colombian National Biodiversity Institute. The goal of this job was to start and use a tissue collection and a molecular lab. This period was very productive and after ten years of work, the collection got more than 10,000 samples of different components of Colombian biodiversity and more than 60 research projects were developed there. During that time continued his academic formation in the same University, but this time in the city of Palmira. He got a Master's Degree in Crops Genetics Resources working with Colombian oaks species and population genetics. After a wonderful experience in Humboldt Institute, he made an interesting change in his life and moved to the agriculture as a biosafety researcher in the National Agriculture and Livestock Institute. That experience was short because and he moved to The United States looking for the "american academic dream" pursuing his PhD.

Permanent email: jdpalacio@gmail.com

This dissertation was typed by the author.